

A Step Beyond Kemperman's Structure Theorem

David J. Gryniewicz^{*†}

February 2, 2008

1 Introduction

Let G be an abelian group, and let A and B be nonempty subsets of G . Their sumset is the set of all pairwise sums, i.e., $A + B = \{a + b \mid a \in A, b \in B\}$. For a set \mathcal{S} of subsets of G , define

$$d^{\subseteq}(A, \mathcal{S}) = \min_{B \in \mathcal{S}} \{d^{\subseteq}(A, B)\},$$

where $d^{\subseteq}(A, B) = |B \setminus A|$, if $A \subseteq B$, and $d^{\subseteq}(A, B) = \infty$ otherwise. Hence $d^{\subseteq}(A, \mathcal{S})$ measures how far away as a subset the set A is from the sets $B \in \mathcal{S}$.

It is the central problem of inverse additive theory to describe the structure of those pairs of subsets A and B with $|A + B|$ small. Such descriptions often prove useful to other related areas of mathematics—a notable example being the use of Freiman's Theorem [6] [25], describing $A \subseteq \mathbb{Z}$ with $|A + A| < c|A|$, to give a more quantitative proof of Szemerédi's Theorem [28] concerning the existence of (4-term) arithmetic progressions in a subset of positive upper density [7].

One of the classical results of inverse additive theory was the complete recursive description given by Kemperman [19] of the 'critical pairs' in an abelian group, i.e., those finite, nonempty subsets A and B such that $|A + B| < |A| + |B|$. Among other applications—including results in graph theory [16] and zero-sum additive theory [8]—Kemperman's Structure Theorem (KST), whose statement we delay until later, yields the descriptions of those subsets of a locally compact abelian group whose Haar measure of the sumset fails to satisfy the triangle inequality [19] [20]. Other applications may also be found in [27] [24].

Unfortunately, KST has not perhaps been appreciated or utilized to its full potential, in part due to the cloud of confusion and misunderstanding stemming from the perceived complexity of the theorem's statement. In fact, several papers have been published with such goals as the simplification of the conclusion of KST [23], the creation of alternative methods for dealing with critical pairs [14] [15], and the clarification of the use of KST in practice [9] [24].

^{*}Universitat Politècnica de Catalunya, Barcelona

[†]This Research was supported in part by the National Science Foundation, as a Math and Physical Sciences Distinguished International Postdoctoral Research Fellow, under grant DMS-0502193.

Concerning the structure of A and B when $A + B$ is small, few precise results besides KST are known for an arbitrary abelian group. In the case $G = \mathbb{Z}$ with $|A| \geq |B|$, then (currently, a few technical restrictions are also needed)

$$|A + B| = |A| + |B| + r \leq |A| + 2|B| - 4 + \epsilon \quad (1)$$

implies that $d^\subseteq(A, \mathcal{AP}_d), d^\subseteq(B, \mathcal{AP}_d) \leq r + 1$ for some d , where \mathcal{AP}_d is the set of arithmetic progressions with difference d , and ϵ equals 0 or 1, depending on a structural condition between A and B [25]. Thus A and B must be large subsets of arithmetic progressions with the same difference.

In the case $G = \mathbb{Z}/p\mathbb{Z}$, where p is prime, then the critical pairs—with $|A| \geq |B| > 1$ and $|A + B| \leq p - 2$ (to avoid three very special degenerate examples)—consist of arithmetic progressions with the same difference [29] [30]. Under some additional restrictions on the cardinality of A , and assuming $A = B$, then a result of Freiman shows $|A + A| \leq 2|A| + r \leq 2.4|A|$ implies that $d^\subseteq(A, \mathcal{AP}_d) \leq r + 1$, for some d [4] [5] [25]; in other words, the above result from \mathbb{Z} holds for $A = B$, by imposing some moderate conditions, in $\mathbb{Z}/p\mathbb{Z}$ as well. The same result is also known for $\mathbb{Z}/p\mathbb{Z}$ under the more general assumption of (1), provided extremely severe conditions are imposed on the cardinalities of A and B [1]. The extent to which the result holds in $\mathbb{Z}/p\mathbb{Z}$ without unnecessary assumptions on the cardinalities is still quite open. Little is known beyond the case $|A + B| = |A| + |B|$, for which Hamidoune and Rødseth established $d^\subseteq(A, \mathcal{AP}_d), d^\subseteq(B, \mathcal{AP}_d) \leq 1$, with only the assumption $|A + B| \leq p - 4$ and the removal of ϵ from (1) [13]; and the case $|A + B| = |A| + |B| + 1$, for which $d^\subseteq(A, \mathcal{AP}_d), d^\subseteq(B, \mathcal{AP}_d) \leq 2$ was shown by Hamidoune, Serra, and Zemor, under similar assumptions with $p > 51$ [12]. Concerning more general abelian groups, for $A \subseteq \mathbb{Z}/m\mathbb{Z}$ with $|A + A| < 2.04|A|$, Deshouillers and Freiman obtained a rough description of A involving computed large constants [3].

In this paper, we move one step beyond KST by completing the description of all subsets A and B that exactly achieve (rather than fail to achieve) the triangle inequality, namely for which $|A + B| = |A| + |B|$. Our main result is Theorem 4.1 (whose statement we also delay until further notation and concepts have been developed), which shows that with a few noted exceptions—all but one in the same vein as the original recursive description of KST—then such A and B must be large subsets of a critical pair; more specifically, there must exist $A' \supseteq A$ and $B' \supseteq B$ such that $|A' + B'| = |A'| + |B'| - 1$, and which contain A and B each with at most one hole, i.e., $|A' \setminus A| \leq 1$ and $|B' \setminus B| \leq 1$. Thus in the case $G = \mathbb{Z}/p\mathbb{Z}$, with p prime, Theorem 4.1 generalizes the prime case completed by Hamidoune and Rødseth, and is the corresponding composite extension of KST. Theorem 4.1 and KST will also yield necessary and sufficient conditions for $|A + B| = |A| + |B|$.

We should remark that Hamidoune, Serra and Zemor very recently established a particular case of Theorem 4.1, under a series of added assumptions, including that $\gcd(|G|, 6) = 1$, that A be a generating subset, that the order of every element of $A \setminus 0$ be at least $|A| + 1$, and that a few smaller technical assumptions also hold [11]—the hypotheses needed for their result, particularly

the assumption on the order of elements, parallel other first-attempt generalizations of additive results, from the prime order case to the more general abelian group setting (see [2] [17] [18] for other such examples).

2 Preliminaries

We will make heavy use of the interpretation of KST given in [9] (or in [10], where the explanations are slightly extended, including the expansion of a minor omission in comment (c.12) of [9]), some of which may also be found in [24]. First we begin by describing many of the important definitions and notation that we will use.

Let G be an abelian group. A subset $A \subseteq G$ is H_a -periodic if A is a union of H_a -cosets, with H_a a subgroup (referred to as the *period*). Note that every set is H_a -periodic with H_a the trivial group. If A is H_a -periodic with H_a a nontrivial subgroup, then A is *periodic*, and otherwise A is *aperiodic*. Note that A being H_a -periodic is equivalent to $A + H_a = A$. Hence if A is H_a -periodic, then so is $A + B$. An H_a -hole in A is an element of $(A + H_a) \setminus A$, and when clear, H_a will be dropped from the notation. A *punctured periodic set*, is a set A such that $A \cup \{\gamma\}$ is periodic for some $\gamma \notin A$, i.e., A contains exactly one H_a -hole for some nontrivial H_a . We remark that a punctured periodic set cannot be periodic (as for instance shown in [9] [10]). We use $\phi_a : G \rightarrow G/H_a$ to denote the natural homomorphism. Note that if A is maximally H_a -periodic (meaning H_a is the maximal subgroup for which A is H_a -periodic, sometimes called the stabilizer), then $\phi_a(A)$ is aperiodic. One of the foundational results of additive theory is the following result of Kneser [21] [22] [19] [25].

Kneser's Theorem. *Let G be an abelian group, and let A_1, A_2, \dots, A_n be a collection of finite, nonempty subsets of G . If $\sum_{i=1}^n A_i$ is maximally H_a -periodic, then*

$$\left| \sum_{i=1}^n \phi_a(A_i) \right| \geq \sum_{i=1}^n |\phi_a(A_i)| - n + 1.$$

Note that if $A + B$ is maximally H_a -periodic and $\rho = |A + H_a| - |A| + |B + H_a| - |B| = d^\subseteq(A, A + H_a) + d^\subseteq(B, B + H_a)$ is the number of holes in A and B , then Kneser's Theorem implies (by multiplying all terms by $|H_a|$) that $|A + B| \geq |A| + |B| - |H_a| + \rho$. Consequently, if either A or B contains a unique element from some H_a -coset, then $|A + B| \geq |A| + |B| - 1$. Also, if $|A + B| \leq |A| + |B| - 1$, then equality holds in the bound from Kneser's Theorem (else $|A + B| \geq |H_a|(|\phi_a(A)| + |\phi_a(B)|) \geq |A| + |B|$).

Given $a_i \in A$ and a subgroup H_a , we use A_{a_i, H_a} to denote $(a_i + H_a) \cap A$, with the H_a dropped from the notation when clear. If $A_{a_i} \neq a_i + H_a$, then A_{a_i} is a *partially filled H_a -coset*. An H_a -decomposition of A is a partition $A_{a_1} \cup \dots \cup A_{a_l}$ of A with $a_i \in A$. The compliment of A is denoted \overline{A} , and we use $\langle A \rangle$ to denote the subgroup generated by A , which, when $0 \in A$, is the smallest subgroup H_a such that $|\phi_a(A)| = 1$.

A *quasi-periodic decomposition* of A with *quasi-period* H_a , where H_a is a nontrivial subgroup, is a partition $A_1 \cup A_0$ of A into two disjoint (each possibly empty) subsets such that A_1 is H_a -periodic or empty, and A_0 is a subset of an H_a -coset. Note every set has a quasi-periodic decomposition with $H_a = G$ and $A_1 = \emptyset$. A set $A \subseteq G$ is *quasi-periodic* if A has a quasi-periodic decomposition $A = A_1 \cup A_0$ with A_1 nonempty. Given a quasi-periodic decomposition $A_1 \cup A_0$ with quasi-period H_a , then A_1 is the *periodic part* of the decomposition, and A_0 is the *aperiodic part* (although it may be periodic if A is periodic). Such a decomposition is *reduced* if A_0 is not quasi-periodic. Note that if A is finite and has a quasi-periodic decomposition $A_1 \cup A_0$ with quasi-period H , then A has a reduced quasi-periodic decomposition $A'_1 \cup A'_0$ with quasi-period $H' \leq H$ and $A'_0 \subseteq A_0$.

An *arithmetic progression* with difference $d \in G \setminus 0$ and length $l \in \mathbb{N}$ is a set of the form $\{a_0, a_0 + d, \dots, a_0 + (l-1)d\}$, where $a_0 \in G$. The terms a_0 and $a_0 + (l-1)d$ are the *end terms* in the progression, with a_0 the first term and $a_0 + (l-1)d$ the last term. Note that an arithmetic progression with difference d is also an arithmetic difference with difference $-d$, with the first and last terms interchanged. A *d-component* of a set A is a maximal arithmetic progression with difference d contained in A . We use $c_d(A)$ to denote the number of aperiodic d -components of A . Note that if a d -component is periodic, then it must be a $\langle d \rangle$ -coset. Hence $c_d(A) = |A + \{0, d\}| - |A|$. A *quasi-progression* with difference d is a set P with a quasi-periodic decomposition $P = P_1 \cup P_0$ with quasi-period $\langle d \rangle$, such that P_0 is an arithmetic progression with difference d . We use $l_d(A)$ to denote the cardinality of the minimal quasi-progression P with difference d that contains A , and $h_d(A) = l_d(A) - |A| = |P \setminus A|$ counts the number of holes in A with respect to such a minimal quasi-progression P .

Assuming $0 \in A \cap B$, for $i \geq 0$ we define the set $N_i(A, B)$ by

- $N_0(A, B) = A$,
- $N_i(A, B) = (A + iB) \setminus (A + (i-1)B)$ for $i \geq 1$,

where $0B = \{0\}$ and $iB = \underbrace{B + \dots + B}_i$ for $i \geq 1$. If the sets A and B are clear, they will often be dropped from the notation. For $x \in G$, let

$$\nu_x(A, B) = (x - A) \cap B.$$

Hence $|\nu_x(A, B)| = |\nu_x(B, A)|$ is the number of representations of $x = a + b$ with $a \in A$ and $b \in B$. For $U \subseteq B$, $i \geq 1$, and N_i nonempty, we use N_i^U to denote the set of all elements $x \in N_i$ such that $\nu_x(A + (i-1)B, B) = U$, and we define $N_i^{\leq U} = \bigcup_{V \subseteq U} N_i^V$. Hence $A + (i-1)B + (B \setminus U) = (A + iB) \setminus N_i^{\leq U}$. In particular, $|N_1^b(A, B)|$ is the number of $a \in A$ with $a + b$ a unique expression element in $A + B$. The sets N_i were first introduced in [26] in connection with small sumsets in $\mathbb{Z}/p\mathbb{Z}$, with p prime, and have since shown themselves to be quite useful [11].

Let \mathcal{P} be the set of periodic subsets of G , let \mathcal{QP} be the set of quasi-periodic subsets of G , let \mathcal{QP}_H be the set of quasi-periodic subsets of G with quasi-period H , let \mathcal{P}_H be the set of H -periodic

subsets of G , let \mathcal{AP} be the set of arithmetic progressions, and let \mathcal{AP}_d be the set of arithmetic progressions with difference d . Note that $\mathcal{P} \subseteq \mathcal{QP}$.

We will need the following basic proposition [25] [19]. Note that if G is finite and $|A| + |B| \geq |G| + r$, then Proposition 2.1 implies $|\nu_g(A, B)| \geq r$ for every $g \in G$.

Proposition 2.1. *Let A and B be nonempty, finite subsets of an abelian group G , and let $r \in \mathbb{Z}$.*

- (i) *If G is finite, and $|A| + |B| \geq |G| + 1$, then $A + B = G$.*
- (ii) *If $|A + B| < |A| + |B| - r$, then $|\nu_g(A, B)| > r$ for every $g \in A + B$.*

We can now give the statement of KST [19].

Kemperman Structure Theorem (KST). *Let A and B be finite, nonempty subsets of an abelian group G . Then $|A + B| = |A| + |B| - 1$, and, moreover, if $A + B$ is periodic then $|\nu_c(A, B)| = 1$ for some c , if and only if there exist quasi-periodic decompositions $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ with common quasi-period H_a , and A_0 and B_0 nonempty, such that:*

- (i) $|\nu_c(\phi_a(A), \phi_a(B))| = 1$, where $c = \phi_a(A_0) + \phi_a(B_0)$
- (ii) $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)| - 1$,
- (iii) $|N_1^b(A, B)| = |N_1^a(B, A)| = 0$ for all $a \in A_1$ and $b \in B_1$, and
- (iv) the pair (A_0, B_0) is of one of the following types (each of which imply $|A_0 + B_0| = |A_0| + |B_0| - 1$):
 - (I) $|A_0| = 1$ or $|B_0| = 1$;
 - (II) A_0 and B_0 are arithmetic progressions with common difference d , where the order of d is at least $|A_0| + |B_0| - 1$, and $|A_0| \geq 2$, $|B_0| \geq 2$ (hence, $A_0 + B_0$ is an arithmetic progression with difference d , while $|\nu_c(A_0, B_0)| = 1$ for exactly two $c \in A_0 + B_0$);
 - (III) $|A_0| + |B_0| = |H_a| + 1$, and precisely one element g_0 satisfies $|\nu_{g_0}(A_0, B_0)| = 1$ (hence, B_0 has the form $B_0 = (g_0 - ((g_1 + H_a) \cap \overline{A_0})) \cup \{g_0 - g_1\}$, where $g_1 \in A_0$);
 - (IV) A_0 is aperiodic, B_0 is of the form $B_0 = g_0 - ((g_1 + H_a) \cap \overline{A_0})$, with $g_1 \in A_0$ (hence, $A_0 + B_0 = (g_0 + H_a) \setminus g_0$), and $|\nu_c(A_0, B_0)| \neq 1$ for all c .

The condition (iii) was not shown in Kemperman's original paper, but can be derived from KST as shown in [9] [10]. Conditions (i) and (ii) imply that the pair $(\phi_a(A), \phi_a(B))$ satisfies the hypothesis of KST. Hence repeated application of KST modulo the quasi-period yields a complete recursive description of the pair (A, B) . This gives rise to a chain of subgroups

$$0 = H_{a_0} < H_a = H_{a_1} < H_{a_2} < \dots < H_{a_n} = G,$$

where $H_{a_i}/H_{a_{i-1}}$ is the quasi-period given by KST after i iterations. Conditions (i) and (iii) ensure 'proper alignment' during this recursive process, namely that $\phi_a(A_0)$ and $\phi_a(B_0)$ will always be contained in the aperiodic part of the mod H_a quasi-periodic decomposition given by KST. Consequently, KST induces partitions (allowing empty parts) $A = A_n \cup A_{n-1} \cup \dots \cup A_1 \cup A_0$ and $B = B_n \cup B_{n-1} \cup \dots \cup B_1 \cup B_0$, such that $A_n = B_n = \emptyset$, A_0 and B_0 are nonempty, and

$$\phi_{a_{j-1}}(A) = \phi_{a_{j-1}}(A_n \cup \dots \cup A_j) \bigcup \phi_{a_{j-1}}(A_{j-1} \cup \dots \cup A_0),$$

and

$$\phi_{a_{j-1}}(B) = \phi_{a_{j-1}}(B_n \cup \dots \cup B_j) \bigcup \phi_{a_{j-1}}(B_{j-1} \cup \dots \cup B_0)$$

are the quasi-periodic decompositions with quasi-period $H_{a_j}/H_{a_{j-1}}$ given by KST after j applications, while

$$A = (A_n \cup \dots \cup A_j) \bigcup (A_{j-1} \cup \dots \cup A_0),$$

and

$$B = (B_n \cup \dots \cup B_j) \bigcup (B_{j-1} \cup \dots \cup B_0)$$

are also quasi-period decompositions of A and B with common quasi-period H_{a_j} .

At first it might appear that the critical pairs with $A + B$ periodic with maximal period H_a , including the cases when $|A + B| < |A| + |B|$, are not fully covered by KST. However these cases are easily reduced to the cases covered by KST. This is because in view of Kneser's Theorem it follows that $|\phi_a(A + B)| = |\phi_a(A)| + |\phi_a(B)| - 1$, with $\phi_a(A + B)$ aperiodic, and $A + B$ containing exactly $|A + B| + |H_a| - |A| - |B|$ holes. Thus KST is used to describe the pair $(\phi_a(A), \phi_a(B))$, and then A and B are obtained from $A + H_a$ and $B + H_a$ by deleting $|A + B| + |H_a| - |A| - |B| \leq |H_a| - 1$ elements. In view of Proposition 2.1, these deleted elements can be placed anywhere in the sets $H_a + A$ and $H_a + B$, with the resulting sets being critical. Thus there is little to say about the location of these holes.

It is natural to wonder if a pair (A, B) could have more than one quasi-periodic decomposition that satisfies KST, in other words, how unique is the representation given by KST. This question is addressed in [9] [10]. We provide a short summary. The type of a critical pair (A, B) is unique, and not dependent on the choice of quasi-periodic decompositions that satisfy KST. If the pair (A, B) has type (II), then there is a unique pair of quasi-periodic decompositions that satisfy KST. The same is true (in view of the added condition (iii)) for type (I). In both cases, the quasi-period H_a may not be unique, but can always be taken to be maximal, subject to the conditions of KST. There is slightly more freedom for types (III) and (IV), but in both cases there is a natural canonical choice. For type (III), the quasi-period may always be taken to be the maximal period of $A + B$. This ensures that type (III) will not occur twice in a row when recursively iterating KST. Type (IV) can only occur in the first iteration of KST (since a type (IV) pair has no unique expression element), and implies that $A + B$ is a punctured H_a -periodic set with $|H_a| \geq 6$. These conditions imply that there is a unique $\gamma \in \overline{A + B}$ such that $A + B \cup \{\gamma\}$ is periodic, and H_a may always be taken to be the maximal period of $A + B \cup \{\gamma\}$. This ensures that type (III) not follow type (IV) when iterating KST. Hence for all types the quasi-period H_a given by KST may always be taken to be maximal, and the resulting quasi-period decompositions that satisfy KST will be referred to as the *Kemperman decompositions*.

Regarding the sequence of possible types, there is some restriction on when type (I) can occur twice in a row, assuming H_a chosen maximally. Namely, if $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ are the Kemperman decompositions of A and B with common quasi-period H_a , and if $\phi_a(A) =$

$\phi_a(A'_1) \cup \phi_a(A'_0)$ and $\phi_a(B) = \phi_a(B'_1) \cup \phi_a(B'_0)$ are the Kemperman decompositions of $\phi_a(A)$ and $\phi_a(B)$, then we cannot have both (A, B) of type (I) with $|A_0| = 1$, and also $(\phi_a(A), \phi_a(B))$ of type (I) with $|\phi_a(A'_0)| = 1$ —this is the main step in the proof of Proposition 2.2 in [9], and likewise for Proposition 5.3 in [10]. Finally, it should also be observed that if A is not quasi-periodic and $\langle -\alpha + A \rangle = G$, with $\alpha \in A$, then $H_a = G$.

We will also need the following two simple propositions that, like the proofs given here, are minor variations on two results from [26].

Proposition 2.2. *Let G be an abelian group, let $X, Y \subseteq G$ be finite and nonempty with $0 \in X \cap Y$, and let $i \geq 1$. If $|\nu_z(X, Y)| \geq t$ for every $z \in X + Y$, then $|\nu_z(X + (i - 1)Y, Y)| \geq t$ for every $z \in X + iY$.*

Proof. It suffices to prove the case $i = 2$, as the other cases follow by repeated application. Let $y, y' \in Y$ and $x \in X$. Since $|\nu_z(X, Y)| \geq t$ for $z \in X + Y$, let $\{x_i + y_i\}_{i=1}^t$ be distinct representations of $x + y$, with $x_i \in X$ and $y_i \in Y$. Hence $\{(x_i + y') + y_i\}_{i=1}^t$ are distinct representations of $x + y + y' \in X + 2Y$ with $x_i + y' \in X + Y$ and $y_i \in Y$, whence $|\nu_z(X + Y, Y)| \geq t$ for every $z \in X + 2Y$. \square

Proposition 2.3. *Let G be an abelian group, let $X, Y \subseteq G$ be finite and nonempty with $0 \in X \cap Y$, and let $i \geq 1$. If $U \subseteq Y$, then $N_{i+1}^U(X, Y) - U \subseteq N_i^{\leq U}(X, Y)$.*

Proof. Let $x \in N_{i+1}^U$. Hence by definition U is the maximal subset of Y such that $x - U \subseteq X + iY$. Furthermore, $x - U \subseteq N_i$, since otherwise $x - u = y$ for some $u \in U \subseteq Y$ and $y \in X + (i - 1)Y$, whence $x = u + y \in X + iY$, contradicting that $x \in N_{i+1}^U \subseteq N_{i+1}$. If $x - U \not\subseteq N_i^{\leq U}$, then for some $u \in U$, it follows that $x - u = z \in N_i$, with $z = u' + x'$ for some $u' \in Y \setminus U$ and $x' \in X + (i - 1)Y$. Thus $x - u' = u + x' \in X + iY$. Hence, since $u' \notin U$, this contradicts the maximality of U . \square

We conclude the section with one last important concept. Given a pair of subsets A and B of an abelian group G , we say that A is *non-extendible* (with respect to B), if

$$A \cup \{a_0\} + B \neq A + B, \text{ for all } a_0 \in \overline{A}. \quad (2)$$

We say A is *extendible* otherwise. We will call the pair (A, B) *non-extendible* if both A and B are non-extendible (with respect to each other), and *extendible* otherwise. Note that in general

$$-B + \overline{A + B} \subseteq \overline{A}, \quad (3)$$

since otherwise $-b + c_0 = a$ for some $b \in B$, $c_0 \in \overline{A + B}$, and $a \in A$, implying $c_0 = a + b \in A + B$, contradicting $c_0 \in \overline{A + B}$. However, A being non-extendible is in fact equivalent to equality holding.

Proposition 2.4. *For subsets A and B of an abelian group, A is non-extendible if and only if*

$$-B + \overline{A + B} = \overline{A}. \quad (4)$$

In particular, the pair (A, B) is non-extendible if and only if both (4) and

$$-A + \overline{A+B} = \overline{B}, \quad (5)$$

hold, and both the pairs $(-A, \overline{A+B})$ and $(-B, \overline{A+B})$ are also non-extendible.

Proof. Suppose A is non-extendible. Let $a_0 \in \overline{A}$. In view of (2), it follows that there exists $b \in B$ and $c_0 \in \overline{A+B}$ such that $a_0 + b = c_0$, whence $a_0 = -b + c_0 \in -B + \overline{A+B}$. Thus $\overline{A} \subseteq -B + \overline{A+B}$, and equality follows in view of (3).

On the other hand, if we suppose (4), then each $a_0 \in \overline{A}$ can be written as $a_0 = -b + c_0$, with $b \in B$ and $c_0 \in \overline{A+B}$, implying $a_0 + b = c_0$ is an element of $A \cup \{a_0\} + B$ not contained in $A + B$, whence (2) follows.

If (A, B) is non-extendible, and $(-A, \overline{A+B})$ is extendible, then in view of (4), (5), and the first part of the proposition, it would follow that either $A + B \neq A + B$ or $-\overline{A+B} + B \neq -\overline{A}$. The former is clearly a contradiction, while in view of (4), the latter contradicts the non-extendibility of A . The same argument applied to the pair $(-B, \overline{A+B})$ shows it to likewise be non-extendible, completing the proof. \square

When G is finite, it is important to note that if $|A+B| = |A| + |B| + r$ with B non-extendible, then proposition 2.4 implies that $|-A + \overline{A+B}| = |-A| + |\overline{A+B}| + r$ as well. Thus a non-extendible pair (A, B) with $|A+B| = |A| + |B| + r$ is part of a triple of non-extendible pairs, all having their sumsets with cardinality exactly r more than the bound given by the triangle inequality. The ideas behind Proposition 2.4 trace their roots back to Vosper [29], and can also be found in the isoperimetric method (since k -fragments are non-extendible) [14] [15].

3 $A + B$ Versus $A - B$

For the proof of our main result, we will need the following basic theorem, which can be viewed as generalization of the bound for Sidon sets. Indeed, if $A = B$, $|T| = 1$ and $k = 1$ (the conditions for A to be a Sidon set, given from the difference set point of view), then the familiar bound $|A+A| \geq \frac{|A|(|A|+1)}{2}$ follows from Theorem 3.1(iii) by noting that $x = |A|$. The bound in Theorem 3.1(i) is a general approximation, whose second half is without constants depending on divisibility, while Theorem 3.1(ii) gives a non-implicit bound for $|T|$, and Theorem 3.1(iii) gives a non-implicit bound for $|A+B|$.

Theorem 3.1 will allow us to conclude $|A+B|$ is large provided $|\nu_x(A, -B)|$ is small for most $x \in A - B$. This will be important as the proof of our main result uses (a variation of) the Dyson e -transform of the pair (A, B) . Namely for $e \in A - B$, where $|A| \geq |B|$, we define $B(e) = (e+B) \cap A$ and $A(e) = (e+B) \cup A$. Then $|A(e)| + |B(e)| = |A| + |B|$ while $A(e) + B(e) \subseteq e + A + B$. Our main strategy will be to apply induction to the pair $(A(e), B(e))$. However, we will encounter problems

with this strategy if $|B(e)|$ is small for all e such that $e + B \not\subseteq A$. In these cases, we will use Theorem 3.1 to show that $|A + B|$ is large instead.

Theorem 3.1. *Let A , B and T be finite subsets of an abelian group G , with $|A| \geq |B| > k \geq 1$ and $|A| \geq |T|$. Let δ be the integer such that $|B|(|A| - |T|) \equiv \delta \pmod{k}$, with $0 \leq \delta < k$, let $M = |T||B|(|B| - k) + (k - 1)|A||B| - \delta(k - \delta)$, let x be the integer, $1 \leq x \leq |A||B|$, such that $M + x \equiv 0 \pmod{|A||B|}$, and let δ_0 be the integer such that $|A||B| + \delta_0 \equiv 0 \pmod{|A + B|}$, with $0 \leq \delta_0 < |A + B|$. If $|\nu_x(A, -B)| \leq k$ for all $x \in G \setminus T$, then the following bounds hold:*

- (i) $|A + B| \geq \frac{|A|^2|B|^2}{M + |A||B|} \geq \frac{|A|^2|B|}{|T|(|B| - k) + k|A|},$
- (ii) $|T| \geq \frac{|A|^2|B|^2 - \delta_0^2 - |A + B|(k|A||B| - \delta_0 - \delta(k - \delta))}{|A + B||B|(|B| - k)} \geq \frac{|A|^2|B|^2 - \delta_0^2 - |A + B|(k|A||B| - \delta_0)}{|A + B||B|(|B| - k)} \geq |A| \frac{|A||B| - k|A + B|}{|A + B|(|B| - k)},$
- (iii) $|A + B| \geq \frac{2|A||B|}{\left\lceil \frac{M + 2|A||B|}{|A||B|} \right\rceil} - \frac{M}{\left\lceil \frac{M + |A||B|}{|A||B|} \right\rceil \left\lceil \frac{M + 2|A||B|}{|A||B|} \right\rceil} = \frac{|A|^2|B|^2(M + 2x)}{(M + |A||B| + x)(M + x)}.$

We first give some basic notions from Graph Theory, in order to put the ideas of the proof of Theorem 3.1 in broader context. For a graph Γ , we use $\bar{\Gamma}$ to denote the complement of Γ , i.e. the graph on the same vertex set with the edges being the non-edges of Γ . Also, $V(\Gamma)$ denotes the vertex set, and $E(\Gamma)$ the edge set. The line graph of Γ , denoted $L(\Gamma)$, is the graph whose vertices are the edges of Γ , with two vertices being adjacent if the corresponding edges in G share a common vertex. Given a pair of nonempty subsets (A, B) of an abelian group G , we can define a graph $\mathcal{M}(A, B)$ whose vertices are the ordered pairs from $A \times B$, with $\{(a, b), (a', b')\}$ an edge precisely when $a + b = a' + b'$. Hence $\mathcal{M}(A, B)$ is a vertex disjoint union of cliques Cl_{x_i} , with each clique Cl_{x_i} corresponding to an element $x_i \in A + B$ such that $|\nu_{x_i}(A, B)| = |V(Cl_{x_i})| = m_i$. Alternatively, $\mathcal{M}(A, B)$ is the complement of the complete multipartite graph corresponding to the sequence m_1, \dots, m_c , where $c = |A + B|$, i.e., $\mathcal{M}(A, B) = \overline{K_{m_1, m_2, \dots, m_c}}$. Letting n_i be the number of cliques Cl_x with $|V(Cl_x)| = i$, we note that

$$|A + B| = \sum_{i \geq 1} n_i,$$

$$|E(\mathcal{M}(A, B))| = \sum_{i \geq 1} n_i \binom{i}{2}. \quad (6)$$

The sumset and difference set of A and B are related via these graphs as follows.

Proposition 3.2. *If A and B are finite, nonempty subsets of an abelian group G , then the map $\varphi : E(\mathcal{M}(A, B)) \rightarrow E(\mathcal{M}(A, -B))$, defined by $\{(a, b), (a', b')\} \mapsto \{(a, -b'), (a', -b)\}$, is a bijection that maps adjacent edges in $L(\mathcal{M}(A, B))$ to distinct components of $\mathcal{M}(A, -B)$. In particular, φ embeds $L(\mathcal{M}(A, B)) \hookrightarrow \overline{L(\mathcal{M}(A, -B))}$, and $|E(\mathcal{M}(A, B))| = |E(\mathcal{M}(A, -B))|$.*

Proof. Since $a + b = a' + b'$ implies $a - b' = a' - b$, it follows that map φ is well defined. Noting that φ is its own inverse, it follows that φ is a bijection. If $a_1 + b_1 = a_2 + b_2 = a_3 + b_3$, with $a_i \in A$

distinct and $b_i \in B$ distinct, then we note that $a_1 - b_2 = a_2 - b_1 = a_2 - b_3 = a_3 - b_2$ would contradict that b_1 and b_3 are distinct. Hence the adjacent edges $\{(a_1, b_1), (a_2, b_2)\}$ and $\{(a_2, b_2), (a_3, b_3)\}$ are mapped to distinct components in $\mathcal{M}(A, -B)$, and thus cannot be adjacent in $L(\mathcal{M}(A, -B))$. \square

It is important to note, by means of a simple extremal argument or discrete derivative, that once $|A + B|$ is fixed, then (6) is minimized by taking all cliques of as near equal a size as possible. Likewise, (6) is maximized in just the opposite way, by taking as many cliques of largest allowed size as possible, followed by as many cliques of the next largest allowed size as possible, \dots , etc.

For the proof of Theorem 3.1, which we now proceed with, we will need only the equality $|E(\mathcal{M}(A, B))| = |E(\mathcal{M}(A, -B))|$.

Proof. Let $|A| = a$, $|B| = b$ and $|T| = t$. Since $|\nu_x(A, -B)| \leq k$ for all $x \in G \setminus T$, since $a \geq b$, and since $a \geq t$, it follows that

$$|E(\mathcal{M}(A, -B))| \leq t \binom{b}{2} + \frac{b(a-t)-\delta}{k} \binom{k}{2} + \binom{\delta}{2} = \frac{tb^2 - tkb + (k-1)ab - \delta(k-\delta)}{2} = \frac{M}{2}. \quad (7)$$

Thus $M \geq 0$, and M is even. Let $l = \left\lceil \frac{ab}{|A+B|} \right\rceil - 1$, and let $c = |A + B|$. Note

$$c = |A + B| = \frac{ab + \delta_0}{l + 1}, \quad (8)$$

where $ab + \delta_0 \equiv 0 \pmod{c}$, and $0 \leq \delta_0 < c$. Since $|E(\mathcal{M}(A, B))|$ will be minimized when $l \leq |\nu_x(A, B)| \leq l + 1$, it follows in view of (8) that

$$|E(\mathcal{M}(A, B))| \geq c \binom{l}{2} + (ab - cl)l = \frac{l(2ab - c(l+1))}{2} = \frac{l(ab - \delta_0)}{2} = \left(\frac{ab + \delta_0}{c} - 1\right) \frac{(ab - \delta_0)}{2}. \quad (9)$$

Considering the above bound as a function of δ_0 , we note that it is quadratic in δ_0 with negative lead coefficient. Thus this quantity will be minimized at a boundary value for δ_0 . Hence comparing the bound evaluated at $\delta_0 = 0$ and $\delta_0 = c - 1$, it follows that

$$|E(\mathcal{M}(A, B))| \geq \left(\frac{ab}{c} - 1\right) \frac{ab}{2}. \quad (10)$$

Since $|E(\mathcal{M}(A, B))| = |E(\mathcal{M}(A, -B))|$ follows from Proposition 3.2, then by comparing (7) and (9), it follows that $c = |A + B|$ must satisfy the bound

$$|A + B| \geq \frac{a^2b^2 - \delta_0^2}{M + ab - \delta_0}, \quad (11)$$

and by comparing (7) and (10), it follows that

$$|A + B| \geq \frac{a^2b^2}{M + ab}. \quad (12)$$

Noting that $M = tb(b-k) + (k-1)ab - \delta(k-\delta) \leq tb(b-k) + (k-1)ab$, it follows that (12) implies (i), and (11) rearranges to yield (ii) (the second two inequalities follow immediately from the first).

Suppose that

$$|A + B| \geq \min \left\{ \frac{|A||B|}{d}, \frac{2|A||B|}{d+1} - \frac{M}{d(d+1)} \right\}, \quad (13)$$

for every integer $d \geq 1$. Comparing the two bounds in the minimum, we see that

$$\frac{2|A||B|}{d+1} - \frac{M}{d(d+1)} \leq \frac{|A||B|}{d}$$

holds for $d \leq \frac{M+ab}{ab}$. Thus, since $M \geq 0$, then applying (13) with $d = \lfloor \frac{M+ab}{ab} \rfloor \geq 1$ yields (iii). So it remains to establish (13).

Note that if $l = 0$, then $|A + B| = |A||B|$ follows from (8), yielding (13). So we may assume $l > 0$. Hence comparing (7) with (9) (expressed in terms of l without δ_0), it follows that

$$|A + B| \geq \frac{2ab}{l+1} - \frac{M}{l(l+1)}. \quad (14)$$

Considering the above bound as a function of l , and computing its discrete derivative, i.e., the bound evaluated at l minus the bound evaluated at $l-1$, for $l \geq 2$, we obtain

$$\frac{2M + 2ab - 2abl}{l(l^2 - 1)}. \quad (15)$$

Noting that (15) is non-negative for $l \leq \frac{M+ab}{ab}$, it follows that the bound in (14) monotonically increases up to $l \leq \frac{M+ab}{ab}$.

Suppose $l > \frac{M+ab}{ab}$. Hence from (8) it follows that

$$\frac{M + ab + 1}{ab} \leq l = \frac{ab + \delta_0}{c} - 1 \leq \frac{ab + c - 1}{c} - 1 < \frac{ab}{c}.$$

Thus, in view of (i), it follows that

$$\frac{a^2b^2}{M + ab} \leq c < \frac{a^2b^2}{M + ab + 1},$$

a contradiction. So $l \leq \frac{M+ab}{ab}$. Consequently, the bound in (14) is monotonically increasing for all possible values of l .

If $l \leq d-1$, then $|A + B| = c \geq \frac{ab}{d}$ follows from (8). Otherwise, $l \geq d$, whence the monotonicity of (14) implies

$$|A + B| \geq \frac{2ab}{d+1} - \frac{M}{d(d+1)}.$$

Hence $|A + B|$ is greater than the minimum of these two bounds, yielding (13), and completing the proof. \square

If $|B| \leq k$, then the restriction $|\nu_x(A, -B)| \leq k$ for all $x \in G \setminus T$ holds trivially. Hence the assumption $|B| > k$ is needed to gain useful information. Likewise, the assumption $|A| \geq |T|$ is not very restrictive, since for $|A| \leq |T|$ it is possible that $|\nu_x(A, -B)| = |B|$ for all $x \in A - B$, in which

case we obtain only the trivial bound for $|E(\mathcal{M}(A, -B))|$ (in such cases, an improved estimate of $|E(\mathcal{M}(A, -B))|$ would yield an improved estimate for $|A + B|$, though we will handle this issue by bounding $|T|$ instead). We remark that the bound given in (iii) is minimized when $x = |A||B|$, which yields (i). However, the estimate $\frac{|A|^2|B|^2}{M+|A||B|}$ can be improved upon, under a variety of circumstances, by using more precise estimates for x . For instance, if $0 \leq |T||B|(|B| - k) - \delta(k - \delta) < |A||B|$, then $x = |A||B| - |T||B|(|B| - k) + \delta(k - \delta)$ follows from the definitions of M and x .

4 A Step Beyond KST

We begin with our main result, Theorem 4.1. Note that the second part of Theorem 4.1 shows that the case with $A + B$ periodic reduces to the case when $A + B$ is aperiodic. In the case when $|G|$ is prime, then the reason for assuming the second part of the bound in (1) was to exclude the cases when $|B|$ is too small to gain exceptional structural information. Indeed, $|A + B| \leq |A||B| \leq |A| + |B| + r$ holds trivially for any pair (A, B) with r sufficiently large compared to both $|A|$ and $|B|$. We note that each of the exceptional cases given by types (V-VII) correspond precisely to those degenerate cases in $\mathbb{Z}/p\mathbb{Z}$ when $\min\{|A|, |B|\}$ is very small, lifted via a quasi-periodic decomposition in precisely the same way the elementary pairs of types (I-IV) were lifted for KST. The last additional exception given by type (VIII) is the first case where a non-quasi-periodic example does not directly correspond to behavior in $\mathbb{Z}/p\mathbb{Z}$. However, the structure of a type (VIII) pair is highly restricted, including $d^\subseteq(A + B, \mathcal{P}), d^\subseteq(\overline{A}, \mathcal{P}), d^\subseteq(\overline{B}, \mathcal{P}) \leq 2$, and cannot occur in a cyclic group. In essence, a type (VIII) pair (A_0, B_0) has both A_0 and B_0 being arithmetic progressions (with common difference) of cosets of a Klein four subgroup H_b , with the exception that all four end terms contain only two (rather than four) elements from the coset, with these elements in the first term of both progressions corresponding to two cosets of the same order two subgroup H_{d_1} of H_b , and these elements in the last term of both progressions corresponding to two cosets of a different order two subgroup H_{d_2} of H_b .

The degenerate cases in $G = \mathbb{Z}/p\mathbb{Z}$ that correspond to when $A + B$ is very close to G (for instance, type (IV) pairs), also have corresponding analogs whenever $d^\subseteq(A + B, \mathcal{P})$ is very small. All this leads one to wonder if $|A + B| = |A| + |B| + r$ and $d^\subseteq(\overline{A}, \mathcal{P}), d^\subseteq(\overline{B}, \mathcal{P}), d^\subseteq(A + B, \mathcal{P}) \geq r + 3$, with equality holding for at most one of the three quantities (note that $|A| \geq d^\subseteq(\overline{A}, \mathcal{P})$, with equality holding when $|G|$ is prime), would always imply there exists a pair (A', B') with $|A' + B'| = |A'| + |B'| - 1$, such that $d^\subseteq(A, A'), d^\subseteq(B, B') \leq r + 1$ (or better yet, a pair (A', B') with $|A' + B'| = |A'| + |B'| + r - 1$, such that $d^\subseteq(A, A'), d^\subseteq(B, B') \leq 1$). If true, this would mean that either one of $\overline{A}, \overline{B}$ or $A + B$ is close to being periodic, or else $|A + B| - |A| - |B|$ bounds how far the pair (A, B) can be from being a critical pair (A', B') .

Theorem 4.1. *Let A and B be finite, nonempty subsets of an abelian group G with $|A + B| =$*

$|A| + |B|$. If $A + B$ is aperiodic, then either there exist $\alpha, \beta \in G$ such that

$$|A \cup \{\alpha\} + B \cup \{\beta\}| = |A \cup \{\alpha\}| + |B \cup \{\beta\}| - 1, \quad (16)$$

or there exist quasi-period decompositions $A = A_1 \cup A_0$ and $B = B_1 \cup B_0$ with common quasi-period H_a , and A_0 and B_0 nonempty, such that

- (i) $|\nu_c(\phi_a(A), \phi_a(B))| = 1$, where $c = \phi_a(A_0) + \phi_a(B_0)$
- (ii) $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)| - 1$,
- (iii) the pair (A_0, B_0) is of one of the following types (each of which imply $|A_0 + B_0| = |A_0| + |B_0|$):
 - (V) $|A_0| = 2$ or $|B_0| = 2$, and $|A_0 + B_0| = |A_0| + |B_0| \leq |H_a| - 2$;
 - (VI) $|A_0| = |B_0| = 3$, $A_0 = b_0 + B_0$ for some $b_0 \in G$, and $|2A_0| > 2|A_0| - 1$;
 - (VII) $|A_0| = 3$ or $|B_0| = 3$, and either $|2A_0| > 2|A_0| - 1$ and $B_0 = b_0 + ((2a_0 - \overline{2A_0}) \cap H_a)$ for some $a_0 \in A_0$ and $b_0 \in B_0$ (if $|A_0| = 3$), or $|2B_0| > 2|B_0| - 1$ and $A_0 = a_0 + ((2b_0 - \overline{2B_0}) \cap H_a)$ for some $a_0 \in A_0$ and $b_0 \in B_0$ (if $|B_0| = 3$), and in both cases $|A_0 + B_0| = |H_a| - 3 \geq 6$;
 - (VIII) there exists a subgroup $H_b \leq H_a$, with $H_b \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, such that both $\phi_b(A_0)$ and $\phi_b(B_0)$ are arithmetic progressions with common difference $\phi_b(d)$, where the order of $\phi_b(d)$ is at least $|\phi_b(A_0)| + |\phi_b(B_0)| - 1$ and $|\phi_b(A_0)|, |\phi_b(B_0)| \geq 2$, such that $a_i + H_b \subseteq A_0$ and $b_i + H_b \subseteq B_0$ for all non-end terms $\phi_b(a_i) \in \phi_b(A_0)$ and $\phi_b(b_i) \in \phi_b(B_0)$, such that $(a_0 + H_b) \cap A_0$ and $(b_0 + H_b) \cap B_0$ are each an H_{d_1} -coset, and such that $(a_l + H_b) \cap A_0$ and $(b_l + H_b) \cap B_0$ are each an H_{d_2} -coset, where H_{d_1} and H_{d_2} are distinct cardinality two subgroups of H_b , where $\phi_b(a_0)$ and $\phi_b(b_0)$ are each the first term in $\phi_b(A_0)$ or $\phi_b(B_0)$, respectively, and where $\phi_b(a_l)$ and $\phi_b(b_l)$ are each the last term in $\phi_b(A_0)$ or $\phi_b(B_0)$, respectively (whence $d^\subseteq(C, \mathcal{Q}\mathcal{P}_{H_{d_i}}) = 1$ and $d^\subseteq(C, \mathcal{P}_{H_{d_i}}) = 2$, for all $C \in \{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$ and $i = 1, 2$).

If $A + B$ is periodic with maximal period H_a , then either (16) holds, or else A and B are H_a -periodic, $\phi_a(A + B)$ is aperiodic, and $|\phi_a(A) + \phi_a(B)| = |\phi_a(A)| + |\phi_a(B)|$.

Observe that (16) implies that a pair (A, B) with $|A + B| = |A| + |B|$ is a large subset of a critical pair (A', B') (with at most one hole in each set). One might also like to know where such holes α and β can be placed in a critical pair (A', B') .

If the pair (A, B) is extendible, then this question is easily answered using KST as follows. The extendibility of (A, B) implies that (16) holds with $A' + B' = A + B$, i.e., that only one element α was deleted, say from A' (i.e., $\beta \in B$), whence there cannot have been any unique expression element of the form $\alpha + b$ in $A' + B'$. If $|\nu_x(A', B')| \geq 2$ for all x , then α can be any element. Thus let $A' = A'_1 \cup A'_0$ and $B' = B'_1 \cup B'_0$ be the Kemperman decompositions with common quasi-period H_a . In view of KST(iii), $\alpha \in A$ can be any element from A'_1 . If we have type (I), then $\alpha \notin A'_0$; if type (II), then α can be any element in A'_0 except the two end terms of the arithmetic progression; if type (III), then $A'_0 = (g_0 - (\overline{B'_0} \cap (g_1 + H_a))) \cup \{g_0 - g_1\}$, where $g_1 \in B'_0$, and α can be any element except $g_0 - g_1$; and if type (IV), then α can be any element of A'_0 .

If the pair (A, B) is non-extendible, then the answer is more complicated, and is provided by the following corollary.

Corollary 4.2. *Let A and B be finite, nonempty subsets of an abelian group G with $|A + B| = |A| + |B|$, and suppose that (16) holds. Let $A' = A \cup \{\alpha\}$ and let $B' = B \cup \{\beta\}$. If (A, B) is non-extendible, then $A' + B'$ cannot be periodic without a unique expression element. Hence we can let $A' = A'_1 \cup A'_0$ and $B' = B'_1 \cup B'_0$ be the Kemperman decompositions with common quasi-period H_a . Furthermore, one of the following must also hold:*

(A) (A', B') has type (II) with both $\alpha \in A'_0$ and $\beta \in B'_0$, both α and β are the second term (by appropriate choice of sign for the progression) in their arithmetic progression with difference d , $|A'_0|, |B'_0| \geq 3$ with equality holding for at most one of the two, and $|\langle d \rangle| \geq |A'_0| + |B'_0| - 1$;

(B) (A, B) is obtainable by applying Proposition 2.4 to a non-extendible type (V) pair (A_0, B_0) from $H_a \leq G$, and inserting this pair (via appropriate translation) into the aperiodic parts of a quasi-periodic decomposition with quasi-period H_a that satisfies KST modulo H_a ;

(C) (A', B') has type (I) such that both the aperiodic parts in the Kemperman decompositions have cardinality one, $\alpha \in A'_1$, $\beta \in B'_1$, and $(\phi_a(A'), \phi_a(B'))$ has type (II) with both $\phi_a(A'_\alpha)$ and $\phi_a(B'_\beta)$ contained in the aperiodic part of the mod H_a Kemperman decomposition, and if $|H_a| > 2$, then both are the second term in their arithmetic progression with $A'_0 + \beta = B'_0 + \alpha = \gamma \notin A + B$, while if $|H_a| = 2$, then either both are the first term in their progression, and $A'_0 + \beta \neq B'_0 + \alpha$ if both progressions have length two, or else both are the second term in their arithmetic progression, $A'_0 + \beta = B'_0 + \alpha = \gamma \notin A + B$, and one of the progressions has at least three terms (in all cases, by appropriate choice of sign for the progressions)

Proof. Since (A, B) is non-extendible and (16) holds, it follows that $(A' + B') \setminus \gamma = A + B$, with $\gamma \in A' + B'$. Furthermore, we may assume that $|N_1^\beta(A', B')| = |N_1^\alpha(B', A')| = 0$, since otherwise the non-extendibility of (A, B) is contradicted by either the critical pair (A', B) or the critical pair (A, B') . Thus γ must have exactly two distinct representations $\alpha + \beta'$ and $\alpha' + \beta$ in $A' + B'$.

Suppose $\alpha \in A'_0$ and $\beta \in B'_0$ (we assume for the moment that $A'_1 \cup A'_0$ and $B'_1 \cup B'_0$ exist, which will be shown to be the case in a following paragraph by similar arguments), and let $A_0 = A'_0 \setminus \alpha$ and $B_0 = B'_0 \setminus \beta$. Hence, since $|N_1^\beta(A', B')| = |N_1^\alpha(B', A')| = 0$, it follows that (A', B') cannot have type (I). If we have type (II), then A'_0 and B'_0 are arithmetic progressions with common difference d , and by choosing an appropriate sign for d , it follows that both α and β must be the second term in their respective progression. Furthermore, $|N_1^\beta(A', B')| = |N_1^\alpha(B', A')| = 0$ implies that $|A'_0|, |B'_0| \geq 3$, and KST implies that $|\langle d \rangle| \geq |A'_0| + |B'_0| - 1$. Additionally, $|A'_0| = |B'_0| = 3$ cannot hold, else $|(A' + B') \setminus (A + B)| \geq 2$, a contradiction. Thus (A) holds.

If we have type (IV), then $|A_0 + B_0| = |H_a| - 2$, whence in view of Proposition 2.4 it follows that (A_0, B_0) is just the result of applying Proposition 2.4 with $G = H_a$ to a non-extendible type (V) pair from the group H_a , and translating appropriately (note type (VI) and (VII) pairs have a similar dual relationship, as do type (I) and (IV) pairs). Thus (B) holds. If we have type (III),

then $d^\subseteq(A + B, \mathcal{P}) = 1$ and $|A_0| + |B_0| = |H_a| - 1$, with both A_0 and B_0 contained in an H_a -coset. Note that there are exactly $|H_a| - |B_0|$ elements $x \in A_0 + H_a$ such that $\gamma \notin x + B_0$. Since $|A_0| < |H_a| - |B_0|$, it follows that there must be at least one such $x \in \overline{A_0} \cap (A_0 + H_a)$, whence adding x contradicts the non-extendibility of (A, B) . Thus type (III) cannot occur for (A', B') in this situation.

Next suppose $A' + B'$ is periodic without a unique expression element. Thus $A + B \cup \{\gamma\}$ is periodic with maximal period H_a , and there are exactly $|H_a| + 1$ holes in A and B . Since $\gamma \notin A + B$, it follows in view of Proposition 2.1 applied with $G = H_a$ that there must exist $a_0 \in A$ and $b_0 \in B$ such that A_{a_0} and B_{b_0} contain at least $|H_a|$ holes with $\gamma \in A_{a_0} + B_{b_0} + H_a$. Thus if the last remaining hole is not also in either A_{a_0} or B_{b_0} , then adding it to either A or B will contradict the non-extendibility of (A, B) . Hence $|A_{a_0}| + |B_{b_0}| = |H_a| - 1$, whence the arguments used in the previous paragraph for type (III) show that the pair (A, B) is extendible. Thus $A' + B'$ cannot be periodic without a unique expression element. Thus, in view of previous cases, we may suppose $\alpha \in A'_1$ or $\beta \in B'_1$, and w.l.o.g. assume the former.

Suppose $|H_a| = 2$. Thus $d^\subseteq(A + B, \mathcal{P}) \leq 2$ and either $|A'_0| = 1$ or $|B'_0| = 1$. Hence (A', B') has type (I), whence $|N_1^\beta(A, B)| = 0$ implies $\beta \in B'_1$ as well. Since γ has exactly two representations in $A' + B'$, since every element of $A'_0 + B'_0$ is a unique expression element, and since $\phi_a(A'_0) + \phi_a(B'_0)$ is a unique expression element, it follows that $\phi_a(A'_0) + \phi_a(B'_0) \neq \phi_a(\gamma)$. Hence, since $\gamma \notin A + B$, it follows that $\phi_a(A'_0) + \phi_a(B'_\beta)$, $\phi_a(B'_0) + \phi_a(A'_\alpha)$, and $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ are the only possible representations of $\phi_a(\gamma) \in \phi_a(A') + \phi_a(B')$.

Suppose $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ is a unique expression element. Hence, since $\phi_a(A'_0) + \phi_a(B'_0)$ is also a unique expression element, and since $\alpha \notin A'_0$ and $\beta \notin B'_0$, it follows that $(\phi_a(A'), \phi_a(B'))$ must have type (II) with (by choosing the appropriate sign for the progression) $\phi_a(A'_0)$ the first term in the arithmetic progression and $\phi_a(A'_\alpha)$ the last term in the arithmetic progression, and the same holding true for $\phi_a(B'_0)$ and $\phi_a(B'_\beta)$. Since $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ is a unique expression element, and since $|H_a| = 2$, it follows that $(A'_\alpha \setminus \alpha) + (B'_\beta \setminus \beta)$ must be missing an element from the coset $\alpha + \beta + H_a$, which must then be γ (since deleting α and β removes only the single element γ from $A' + B'$). Hence $\phi_a(A'_\alpha) + \phi_a(B'_\beta) = \phi_a(\gamma)$. Thus, since $\phi_a(A'_\alpha) + \phi_a(B'_\beta) = \phi_a(\gamma)$ is a unique expression element, it follows that $|A'_0| = |B'_0| = 1$, else adding the other element from the H_a -coset either to A'_0 or B'_0 will contradict the non-extendibility of (A, B) . Furthermore, if there are only two terms in each progression, then it follows that $(A'_0 + (B'_\beta \setminus \beta)) \cup (B'_0 + (A'_\alpha \setminus \alpha))$ is an H_a -coset, else $A + B$ will be missing an element besides γ from the coset $A'_0 + B'_\beta + H_a = B'_0 + A'_\alpha + H_a$, contradicting that $(A' \setminus \alpha) + (B' \setminus \beta) = (A' + B') \setminus \gamma$. Hence $A'_0 + \beta \neq B'_0 + \alpha$ in this case. Thus (C) holds. So next assume $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ is not a unique expression element.

If $\phi_a(A'_\alpha) + \phi_a(B'_\beta) = \phi_a(\gamma)$, then there will be two representations of γ contained in $A'_\alpha + B'_\beta$ (since $\alpha \in A'_1$ and $\beta \in B'_1$), whence $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ not a unique expression element implies that γ must have at least one more representation in $A' + B'$ (since $\phi_a(A'_0) + \phi_a(B'_0)$ is a unique expression element, and hence not equal to $\phi_a(A'_\alpha) + \phi_a(B'_\beta) = \phi_a(\gamma)$, and since all other pairs

$\phi_a(b_1) \in \phi_a(A')$ and $\phi_a(b_2) \in \phi_a(B')$ have either $b_1 + H_a \subseteq A'$ or $b_2 + H_a \subseteq B'$, a contradiction. Therefore $\phi_a(A'_\alpha) + \phi_a(B'_\beta) \neq \phi_a(\gamma)$.

Thus, since γ has exactly two representations in $A' + B'$ given by $\alpha + \beta'$ and $\alpha' + \beta$, and since $\phi_a(A'_0) + \phi_a(B'_\beta)$, $\phi_a(B'_0) + \phi_a(A'_\alpha)$, and $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ are the only possible representations of $\phi_a(\gamma) \in \phi_a(A') + \phi_a(B')$, it follows that $\phi_a(A'_0) + \phi_a(B'_\beta) = \phi_a(B'_0) + \phi_a(A'_\alpha) = \phi_a(\gamma)$ is an element of $\phi_a(A') + \phi_a(B')$ with exactly two representations, and that $|A'_0| = |B'_0| = 1$. From KST it follows, as remarked in Section 2, that if (A', B') has type (I) with $|A'_0| = 1$, then $(\phi_a(A'), \phi_a(B'))$ cannot have type (I) with the aperiodic part of $\phi_a(A')$ having cardinality one.

Let H_d/H_a be the quasi-period from KST applied modulo H_a to $(\phi_a(A'), \phi_a(B'))$. Hence, since $|\phi_a(A')_{\phi_a(b), H_d/H_a}| \geq 2$ for all $b \in A'$ (in view of the previous paragraph), it follows that

$$|(\phi_a(\beta) + H_d/H_a) \cap (\phi_a(\gamma) - \phi_a(A'))| \geq 2.$$

Thus, since $\phi_a(A'_0) + \phi_a(B'_\beta) = \phi_a(B'_0) + \phi_a(A'_\alpha) = \phi_a(\gamma)$ is an element of $\phi_a(A') + \phi_a(B')$ with exactly two representations, it follows that either $\phi_a(B'_\beta)$ and $\phi_a(B'_0)$ are contained in the same H_d/H_a -coset, or else $\phi_a(B'_\beta)$ must be contained in the H_d/H_a -coset corresponding to the aperiodic part of the Kemperman decomposition. However, since $\phi_a(B'_0)$ must be contained in the H_d/H_a -coset corresponding to the aperiodic part of the Kemperman decomposition (a consequence of KST(i)(iii)), it follows that $\phi_a(B'_\beta)$ is also contained in the aperiodic part the Kemperman decomposition modulo H_a in the former case as well. Likewise for $\phi_a(A'_\alpha)$.

If $(\phi_a(A'), \phi_a(B'))$ has type (II), then (by appropriate choice of sign), $\phi_a(A'_0)$ is the first term in the arithmetic progression, and $\phi_a(A'_\alpha)$ is the second term, with the same true of $\phi_a(B'_0)$ and $\phi_a(B'_\beta)$ (in view of the previous paragraph, and since $\phi_a(A'_\alpha) + \phi_a(B'_0) = \phi_a(A'_0) + \phi_a(B'_\beta)$ is an element with exactly two distinct representations in $\phi_a(A') + \phi_a(B')$). Furthermore, $\phi_a(A'_\alpha) + \phi_a(B'_\beta)$ not a unique expression element implies that $\phi_a(A'_\alpha)$ and $\phi_a(B'_\beta)$ cannot also both be end terms, whence there must be at least three terms in one of the progressions. Additionally, since $\phi_a(B'_0) + \phi_a(A'_\alpha) = \phi_a(A'_0) + \phi_a(B'_\beta) = \phi_a(\gamma)$, it follows that $\gamma \notin (A'_0 + (B'_\beta \setminus \beta)) \cup (B'_0 + (A'_\alpha \setminus \alpha))$, whence $A'_0 + \beta = B'_0 + \alpha = \gamma$. Thus (C) holds.

Note $(\phi_a(A'), \phi_a(B'))$ cannot have type (IV), nor, as remarked earlier, type (I). If instead $(\phi_a(A'), \phi_a(B'))$ has type (III), then $\phi_a(A'_0) + \phi_a(B'_\beta) = \phi_a(B'_0) + \phi_a(A'_\alpha)$ being an element of $\phi_a(A') + \phi_a(B')$ with exactly two representations would imply $\phi_a(A'_0) + \phi_a(B'_\beta)$ was a unique expression element in $\phi_a(A') + \phi_a(B' \setminus B'_0)$. Since a type (III) pair has exactly one unique expression element, it follows that $|\phi_a(A'_0)|, |\phi_a(B'_0)| \geq 3$. Hence, since $\phi_a(A'_0) + \phi_a(B'_0)$ is the unique expression element in $\phi_a(A') + \phi_a(B')$, it follows from KST that $(\phi_a(A'), \phi_a(B' \setminus B'_0))$ either has type (IV), whence there cannot be any unique expression element, a contradiction, or else type (II). In the latter case, then $\phi_a(A'_0) + \phi_a(B'_\beta)$ a unique expression element in $(\phi_a(A') + \phi_a(B' \setminus B'_0))$ implies that $\phi_a(A'_0)$ is an end term of the arithmetic progression, whence $\phi_a(A'_0) + \phi_a(B'_0)$ a unique expression element in $\phi_a(A') + \phi_a(B')$ implies that $\phi_a(B'_0)$ must follow directly after/before the end term of the progression $\phi_a(B' \setminus B'_0)$, whence $(\phi_a(A'), \phi_a(B'))$ has type (II), a contradiction. Thus type (III)

cannot occur for $(\phi_a(A'), \phi_a(B'))$ in this situation.

Finally, assume instead $|H_a| > 2$. Since $\alpha \in A'_1$, then from Proposition 2.1, it follows that $(A'_\alpha \setminus \alpha) + B_{b_i}$ will be H_a -periodic for all $|B_{b_i}| \geq 2$. Thus $\gamma \notin A + B$ and $|H_a| > 2$ imply that $\phi_a(A'_\alpha) + \phi_a(B'_0) = \phi_a(\gamma)$, and that either $|B'_0| = 1$ or else $|B'_0| = 2$ and $\beta \in B'_0$. However, in the latter case, $|B'_0| = 2$ implies that type (I) or (II) must hold (since type (III) and (IV) both imply that each of the aperiodic parts of the Kemperman decompositions have cardinality at least three), whence $|B'_0| = 2$ further implies that $|N_1^b(A, B)| > 0$ for all $b \in B'_0$, contradicting that $|N_1^\beta(A, B)| = 0$. Therefore we can assume $|B_0| = 1$, whence we must have type (I). Hence $|N_1^\beta(A, B)| = 0$ implies that $\beta \in B_1$ also. Applying the same arguments with the roles of A and B reversed, it follows that $|B_0| = |A_0| = 1$ with $\phi_a(A'_0) + \phi_a(B'_\beta) = \phi_a(B'_0) + \phi_a(A'_\alpha) = \phi_a(\gamma)$ an element of $\phi_a(A' + B')$ with exactly two representations (since $|\nu_\gamma(A', B')| = 2$), whence the arguments from the previous three paragraphs apply for determining the structure modulo H_a , with the exception that the arithmetic progressions from KST are allowed to both have length two, yielding (C), and completing the proof. \square

We remark that the sufficiency of the examples (A-C) is easily checked, as is the sufficiency for types (V-VIII) (in view of Proposition 2.1, Lemma 4.6, and KST). Thus together, along with the description provided for extendible pairs (A, B) , they may be taken as necessary and sufficient conditions for $|A + B| = |A| + |B|$. Simple consequences of the above description and Theorem 4.1 are the following two immediate corollaries.

Corollary 4.3. *Let A and B be finite, nonempty, and non-extendible subsets of an abelian group G such that $|A + B| = |A| + |B|$. If A and B are non-quasi-periodic, generating subsets such that $|A|, |B|, |\overline{A + B}| \geq 3$, with equality holding for at most one of the three, then*

$$d^\subseteq(A + B, \mathcal{S}) = d^\subseteq(\overline{A + B}, \mathcal{S}) = 1, \text{ and}$$

$$d^\subseteq(B, \mathcal{S}) = d^\subseteq(\overline{B}, \mathcal{S}) = d^\subseteq(A, \mathcal{S}) = d^\subseteq(\overline{A}, \mathcal{S}) = 1,$$

where either $\mathcal{S} = \mathcal{QP}_H$, for a nontrivial, proper subgroup H , or else $\mathcal{S} = \mathcal{AP}_d$, for a nonzero $d \in G$.

Corollary 4.4. *Let A and B be finite, nonempty subsets of an abelian group G such that $|A + B| = |A| + |B|$ and $d^\subseteq(A, \mathcal{QP}), d^\subseteq(B, \mathcal{QP}) \geq 2$. If $d^\subseteq(\overline{A}, \mathcal{P}), d^\subseteq(\overline{B}, \mathcal{P}), d^\subseteq(A + B, \mathcal{P}) \geq 3$, with equality holding for at most one of the three, then for some nonzero $d \in G$,*

$$d^\subseteq(A + B, \mathcal{AP}_d), d^\subseteq(A, \mathcal{AP}_d), d^\subseteq(B, \mathcal{AP}_d) \leq 1.$$

The proof of Theorem 4.1 is heavily reliant upon a series of reductions to simpler cases. An important step in the proof will be to show that it suffices to prove Theorem 4.1 when both A and B are non-quasi-periodic, generating subsets. This will be accomplished principally through the following three lemmas.

Lemma 4.5. *Let A and B be finite, nonempty subsets of an abelian group G with $|A| \geq 3$, $|A+B| = |A| + |B|$, and $0 \in A$, and let $H_a = \langle A \rangle$. If $A+B$ is aperiodic and B is non-extendible, then B has a quasi-periodic decomposition with quasi-period H_a .*

Proof. Let B' be the maximal subset of B that is H_a -periodic, and let $B \setminus B' = B_{b_1} \cup \dots \cup B_{b_l}$ be an H_a -coset decomposition of $B \setminus B'$. From the maximality of B' it follows that no B_{b_i} is H_a -periodic. Hence, since B is non-extendible, it follows that

$$A + B_{b_i} \neq b_i + H_a, \quad (17)$$

for all i . Since $|A+B| = |A| + |B|$, since B' is H_a -periodic, and since $|\phi_a(A)| = 1$, it follows that

$$|A + (B \setminus B')| = |A| + |B \setminus B'|. \quad (18)$$

If the lemma is false, then $l \geq 2$. Thus

$$|A + (B \setminus B')| = |\cup_{i=1}^l A + B_{b_i}| \geq |A + B_{b_1}| + |A + B_{b_2}| + |B \setminus \{B' \cup B_{b_1} \cup B_{b_2}\}|,$$

implying in view of (18) that

$$|A + B_{b_1}| + |A + B_{b_2}| = |A + (B_{b_1} \cup B_{b_2})| \leq |A| + |B_{b_1}| + |B_{b_2}|. \quad (19)$$

In view of Kneser's Theorem, it follows that

$$|A + B_{b_i}| \geq |A| + |B_{b_i}| - |H_{a_i}| + \rho_i, \quad (20)$$

where $A + B_{b_i}$ is maximally H_{a_i} -periodic and ρ_i is the number of H_{a_i} -holes in A and B_{b_i} . Since $H_a = \langle A \rangle$, it follows in view of (17) that $|\phi_{a_i}(A)| \geq 2$, whence

$$|A| \geq |\phi_{a_i}(A)| |H_{a_i}| - \rho_i \geq 2|H_{a_i}| - \rho_i. \quad (21)$$

Combining (20) and (19), and w.l.o.g. assuming $|H_{a_1}| \geq |H_{a_2}|$, it follows that

$$|A| \leq |H_{a_1}| + |H_{a_2}| - \rho_1 - \rho_2 \leq 2|H_{a_1}| - \rho_1 - \rho_2. \quad (22)$$

Thus in view of (21) applied with $i = 1$, it follows that $\rho_2 = 0$. Hence, since A is aperiodic (else $A+B$ is periodic), it follows that $|H_{a_2}| = 1$, whence (22) and (21) imply $|H_{a_1}| = 1$ also. Thus (22) implies $|A| \leq 2$, a final contradiction. \square

Lemma 4.6. *Let A and B be finite, nonempty subsets of an abelian group G with $A+B$ aperiodic and (A, B) non-extendible, and let $A = A_1 \cup A_0$ be a quasi-periodic decomposition with A_1 nonempty and periodic with maximal period H_a . If $|A+B| = |A| + |B|$, then B has a quasi-periodic decomposition $B = B_1 \cup B_0$ with quasi-period H_a , such that:*

- (i) $|\nu_c(\phi_a(A), \phi_a(B))| = 1$, where $c = \phi_a(A_0 + B_0)$
- (ii) $|\phi_a(A+B)| = |\phi_a(A)| + |\phi_a(B)| - 1$,
- (iii) $|A_0 + B_0| = |A_0| + |B_0|$.

Proof. Let B' be the maximal subset of B that is H_a -periodic, and let $B \setminus B' = B_{b_1} \cup \dots \cup B_{b_l}$ be an H_a -coset decomposition of $B \setminus B'$. From the maximality of B' it follows that no B_{b_i} is H_a -periodic. Hence, since B is non-extendible, it follows, for all i , that

$$A_0 + B_{b_i} \neq b_i + H_a. \quad (23)$$

If $|\phi_a(A_1) + \phi_a(B)| < |\phi_a(A_1)| + |\phi_a(B)| - 1$, then from Kneser's Theorem it follows that $\phi_a(A_1) + \phi_a(B)$ is periodic, contradicting either the maximality of H_a for A_1 or the fact that A is non-extendible. Therefore

$$|\phi_a(A_1) + \phi_a(B)| \geq |\phi_a(A_1)| + |\phi_a(B)| - 1. \quad (24)$$

Since $A + B$ is aperiodic, it follows that $|H_a||\phi_a(B)| > |B|$ and that $|A + B| \geq |A_1 + B| + |A_0|$. Hence if (24) is strict, then

$$|A + B| \geq |H_a|(|\phi_a(A_1)| + |\phi_a(B)|) + |A_0| \geq |A| + |B| + 1, \quad (25)$$

a contradiction. Therefore we can assume

$$|\phi_a(A_1) + \phi_a(B)| = |\phi_a(A_1)| + |\phi_a(B)| - 1. \quad (26)$$

Likewise, if $A_0 + B' \not\subseteq A_1 + B$, then $|A + B| \geq |A_1 + B| + |H_a| + |A_0|$, whence (25) again follows, a contradiction. So we may assume otherwise.

From the non-extendibility of B , and in view of (23), it follows that $A_0 + B_{b_i}$ is disjoint from $A_1 + B$ for all i . Thus $\phi_a(A_0) + \phi_a(B_{b_i})$ is a unique expression element for each i . Hence, since $A_0 + B' \subseteq A_1 + B$, it follows in view of (26) that

$$|\phi_a(A + B)| = |\phi_a(A)| + |\phi_a(B)| - 2 + l. \quad (27)$$

Note (i) and (ii) along with $|A + B| = |A| + |B|$ force (iii) by a simple counting argument. Consequently, the proof will be complete if $l = 1$. So assume $l \geq 2$.

From Kneser's Theorem applied via translation with group H_a , it follows that

$$|A_0 + B_{b_i}| \geq |A_0| + |B_{b_i}| - |H_{a_i}|, \quad (28)$$

where $A_0 + B_{b_i}$ is maximally H_{a_i} -periodic, and $H_{a_i} \leq H_a$. In view of (26), and since each $A_0 + B_{b_i}$ is disjoint from $A_1 + B$, it follows that

$$\begin{aligned} |A + B| &\geq (|\phi_a(A_1)| + |\phi_a(B')| + l - 1)|H_a| + \sum_{i=1}^l |A_0 + B_{b_i}| = \\ &|A| + |B'| - |A_0| + (l - 1)|H_a| + \sum_{i=1}^l |A_0 + B_{b_i}| \geq \\ &|A| + |B| - |A_0| - |B_{b_1}| - |B_{b_l}| + (l - 1)|H_a| + |A_0 + B_{b_1}| + |A_0 + B_{b_2}|. \end{aligned} \quad (29)$$

In view of (23), it follows that $|H_{a_i}| \leq \frac{1}{2}|H_a|$. Thus (28), (29) and $l \geq 2$ together imply that

$$|A + B| \geq |A| + |B| + |A_0|,$$

whence $|A_0| = 0$ and $A = A_1$, contradicting that $A + B$, and thus A , is aperiodic. \square

Lemma 4.7. *Let A and B be finite, nonempty subsets of an abelian group G with $|A+B| = |A|+|B|$, $0 \in A \cap B$, and $|A|, |B| \geq 3$. If $A + B$ is aperiodic, (A, B) is non-extendible, and neither A nor B is quasi-periodic, then $\langle A \rangle = \langle B \rangle$.*

Proof. Since B is not quasi-periodic, it follows from Lemma 4.5 that $\langle B \rangle \leq \langle A \rangle$. Likewise $\langle A \rangle \leq \langle B \rangle$, whence $\langle A \rangle = \langle B \rangle$. \square

The following lemma essentially shows that it is sufficient for A to be a non-quasi-periodic, generating subset in order that this be true (at least when G is finite) of every $C \in \{A, B, A + B, \overline{B}, \overline{A}, \overline{A+B}\}$. In the proof, via various means, we will often reduce the case $A + B = C$ to a case $A' + B' = C'$, where at least one of A' , B' and C' is a set from $\pm\{A, B, A + B, \overline{B}, \overline{A}, \overline{A+B}\}$, and then employ an induction hypothesis to $A' + B' = C'$, or the like. However, since we will want to stay restricted to the class of non-quasi-periodic, generating subsets, the following lemma, along with Proposition 2.4, will allow us to transfer these assumptions from $A + B = C$ to $A' + B' = C'$.

Lemma 4.8. *Let A and B be finite, nonempty subsets of an abelian group G with $|A+B| = |A|+|B|$, $0 \in A \cap B$, $|A|, |B|, |\overline{A+B}| \geq 3$, $A + B$ aperiodic and (A, B) non-extendible. If $\langle A \rangle = G$ and A is not quasi-periodic, then B is not quasi-periodic and $\langle B \rangle = G$. Furthermore, if G is finite, then neither $A + B$ nor $\overline{A+B}$ is quasi-periodic, and $\langle -\gamma + \overline{A+B} \rangle = G$, where $\gamma \in \overline{A+B}$.*

Proof. If B is quasi-periodic with quasi-period H_a , then since A is not quasi-periodic, it follows in view of Lemma 4.6 that $G = \langle A \rangle \leq H_a$, implying $H_a = G$. Hence $B = G$, contradicting that $A + B$ is aperiodic. Therefore we can assume B is not quasi-periodic. Hence from Lemma 4.7 it follows that $G = \langle A \rangle = \langle B \rangle$.

Now assume G is finite. Since (A, B) is non-extendible, it follows in view of Proposition 2.4 that $-A + \overline{A+B} = \overline{B}$, with $(-A, \overline{A+B})$ non-extendible. Since $A+B$ is aperiodic, it follows that B , and thus \overline{B} , is aperiodic. Hence from the result of the previous paragraph applied to $-A$ and $\overline{A+B}$, it follows that $\overline{A+B}$ is not quasi-periodic and $\langle -\gamma + \overline{A+B} \rangle = \langle A \rangle = G$, where $\gamma \in \overline{A+B}$. If $A+B$ is quasi-periodic with quasi-period H_a , then it follows, in view of $\overline{A+B}$ not quasi-periodic, that $G = \langle -\gamma + \overline{A+B} \rangle \leq H_a$. Hence $H_a = G$, contradicting that $A + B$ is aperiodic, and completing the proof. \square

Next, we prove the following simple proposition that will be needed in several places.

Proposition 4.9. *Let A and B be finite subsets of an abelian group, and let $C \subseteq A + B$. Suppose $|A|, |B|, |C| \geq 2$. If $(a + B) \cap C$ and $(b + A) \cap C$ are both nonempty for every $a \in A$ and $b \in B$,*

then either (a) there exist distinct $a, a' \in A$, distinct $b, b' \in B$ and distinct $c, c' \in C$ such that $a + b = c$ and $a' + b' = c'$, or else (b) $|A| = |B| = |C| = 2$, and there exists $a \in A$ and $b \in B$ such that $a + B = b + A = C$.

Proof. Since $C \subseteq A + B$, let $a_1 + b_1 = c_1$ with $a_1 \in A$, $b_1 \in B$ and $c_1 \in C$. Since $|C| \geq 2$, let $c_2 \in C$ be distinct from c_1 . Since $C \subseteq A + B$, it follows that $c_2 = a + b$ for some $a \in A$ and $b \in B$. If $a \neq a_1$ and $b \neq b_1$, then the proof is complete. Otherwise, we may w.l.o.g. by symmetry assume $c_2 = a_2 + b_1$, with $a_2 \neq a_1$ (since $c_2 \neq c_1$). Since $|B| \geq 2$, let $b_2 \in B$ with $b_2 \neq b_1$. Since $(b_2 + A) \cap C$ is nonempty, it follows that $a + b_2 = c$ for some $a \in A$ and $c \in C$. If $c \notin \{c_1, c_2\}$, then (a) is satisfied with $a + b_2 = c$ and either $a_1 + b_1 = c_1$ (if $a \neq a_1$) or $a_2 + b_1 = c_2$ (if $a \neq a_2$). Thus we can w.l.o.g. by symmetry assume $c = c_2$. If $a \notin \{a_1, a_2\}$, then (a) is satisfied with $a + b_2 = c_2$ and $a_1 + b_1 = c_1$. Thus, since $a_2 + b_1 = c_2$ (implying $a_2 + b_2 \neq c_2$), it follows that we can assume $a = a_1$. Hence $a_1 + b_2 = c_2$, $a_1 + \{b_1, b_2\} = \{c_1, c_2\}$ and $b_1 + \{a_1, a_2\} = \{c_1, c_2\}$.

Suppose there exists $a \in A \setminus \{a_1, a_2\}$. Hence, since $(a + B) \cap C$ is nonempty, it follows that $a + b = c$ for some $b \in B$ and $c \in C$. If $c \notin \{c_1, c_2\}$, then (a) is satisfied with $a + b = c$ and either $a_1 + b_1 = c_1$ (if $b \neq b_1$) or $a_1 + b_2 = c_2$ (if $b \neq b_2$). Therefore we can assume $c \in \{c_1, c_2\}$. If $b \notin \{b_1, b_2\}$, then (a) is satisfied with $a + b = c$ and either $a_1 + b_1 = c_1$ (if $c \neq c_1$) or $a_1 + b_2 = c_2$ (if $c \neq c_2$). Therefore, since $a_1 + b_1 = c_1$ and $a_2 + b_1 = c_2$ (implying $a + b_1 \notin \{c_1, c_2\}$), and since $a_1 + b_2 = c_2$ (implying $a + b_2 \neq c_2$), it follows that $b = b_2$ and $c = c_1$. Thus (a) is satisfied with $a_2 + b_1 = c_2$ and $a + b_2 = c_1$. So we can assume $|A| = 2$. Likewise, the same argument applied to B shows that $|B| = 2$.

Suppose $c \in C \setminus \{c_1, c_2\}$. Since $C \subseteq A + B$, it follows that $a + b = c$ for some $a \in A$ and $b \in B$. Since $|A| = |B| = 2$, and since $a_1 + b_1 = c_1$, $a_1 + b_2 = c_2$, and $a_2 + b_1 = c_2$, it follows that $c = a_2 + b_2$, whence (a) is satisfied with $a_2 + b_2 = c$ and $a_1 + b_1 = c_1$. So we can assume $|C| = 2$. Hence in view of the conclusion of the first paragraph, it follows that $a_1 + B = a_1 + \{b_1, b_2\} = \{c_1, c_2\} = C$ and $b_1 + A = b_1 + \{a_1, a_2\} = \{c_1, c_2\} = C$, whence (b) holds. \square

In the proof, once we have established that Theorem 4.1 holds for $A' + B' = C'$, we will want to transfer the resulting structure back to $A + B = C$. Since one of A' , B' and C' will be a set from $\pm\{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$, our strategy will be to use this common linking set (along with Proposition 2.4) as the means of transferring the structural information. However, to accomplish this, we will need to know that having the unpaired structural information for only the set A , is enough to conclude that Theorem 4.1 holds for the pair that includes A . The following two lemmas accomplish this in the case when a non-quasi-periodic, generating subset A is close to being quasi-periodic. Note that we have begun assuming G finite in the hypotheses of the lemmas. This is done to make use of the set $\overline{A + B}$ via Proposition 2.4. However, since we will be able to later show that the case G infinite follows from the case G finite, this will not hinder the proof.

Lemma 4.10. *Let A and B be nonempty subsets of a finite abelian group G with $|A + B| = |A| + |B|$, $0 \in A \cap B$, $|A|, |B|, d^\subseteq(A + B, \mathcal{P}) \geq 3$, (A, B) non-extendible, $\langle A \rangle = G$ and A not quasi-periodic.*

If $A = A' \cup A_2 \cup A_1$ with A' a nonempty periodic subset with maximal period H_a , $A_1 \subseteq a_1 + H_a$, and $A_2 \subseteq a_2 + H_a$, for some $a_i \in G$, then $d^\subseteq(C, \mathcal{QP}) = 1$ for all $C \in \{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$ and (16) holds.

Proof. Let B' be the maximal subset of B that is H_a -periodic, and let $B \setminus B' = B_{b_1} \cup \dots \cup B_{b_l}$ be an H_a -coset decomposition of $B \setminus B'$. From the maximality of B' it follows that no B_{b_i} is H_a -periodic. From Lemma 4.8 it follows that $\langle B \rangle = G$ and that B is not quasi-periodic. Hence $l \geq 2$, as otherwise $H_a = G$ implying $A = G$, which contradicts $A + B$ aperiodic. Since A is not quasi-periodic, it follows that A_1 and A_2 are both partially filled H_a -cosets that are nonempty and disjoint modulo H_a .

If $|\phi_a(A') + \phi_a(B)| < |\phi_a(A')| + |\phi_a(B)| - 1$, then from Kneser's Theorem it follows that $\phi_a(A') + \phi_a(B)$ is periodic, contradicting either the maximality of H_a for A' or the fact that A is non-extendible. Therefore

$$|\phi_a(A') + \phi_a(B)| \geq |\phi_a(A')| + |\phi_a(B)| - 1. \quad (30)$$

Let C' be the maximal subset of $A + B$ that is H_a -periodic, and let $(A + B) \setminus C' = C_{c_1} \cup C_{c_2} \cup \dots \cup C_{c_r}$ be an H_a -coset decomposition. In view of Lemma 4.8 it follows that $A + B$ is not quasi-periodic. Thus $r \geq 2$ (as $A' \neq \emptyset$ implies $C' \neq \emptyset$). Note that $A' + B \subseteq C'$. Hence from (30) it follows that

$$\begin{aligned} |A + B| &\geq (|\phi_a(A')| + |\phi_a(B')| + l - 1)|H_a| + \sum_{i=1}^r |C_{c_i}| \geq \\ &|A| + |B| - (|A_1| + |A_2| + \sum_{i=1}^l |B_{b_i}|) + (l - 1)|H_a| + \sum_{i=1}^r |C_{c_i}|. \end{aligned} \quad (31)$$

Since the C_{c_i} are partially filled H_a -cosets, it follows that $\phi_a(C \setminus C') \subseteq \phi_a(A_1 \cup A_2) + \phi_a(B \setminus B')$. From the non-extendibility of A , and since each A_i is a partially filled H_a -coset, it follows that each a_i has at least one b_j such that $\phi_a(a_i) + \phi_a(b_j) \in \phi_a(C \setminus C')$. In view of the non-extendibility of B , and since each B_{b_i} is a partially filled H_a -coset, it follows that same holds true for each b_i . Consequently, it follows in view of Proposition 4.9 that either there exist distinct b_{i_1} and b_{i_2} , and distinct $c_{i'_1}$ and $c_{i'_2}$, such that $\phi_a(a_1 + b_{i_1}) = \phi_a(c_{i'_1})$ and $\phi_a(a_2 + b_{i_2}) = \phi_a(c_{i'_2})$, or else $l = r = 2$ and w.l.o.g. $\phi_a(a_1 + \{b_1, b_2\}) = \{\phi_a(c_1), \phi_a(c_2)\}$ and $\phi_a(b_1 + \{a_1, a_2\}) = \{\phi_a(c_1), \phi_a(c_2)\}$. We handle these cases separately.

Case 1: First assume there exist distinct b_{i_1} and b_{i_2} , and distinct $c_{i'_1}$ and $c_{i'_2}$, such that $\phi_a(a_1 + b_{i_1}) = \phi_a(c_{i'_1})$ and $\phi_a(a_2 + b_{i_2}) = \phi_a(c_{i'_2})$. Since the C_{c_i} are partially filled, it follows from Kneser's Theorem that $|A_1 + B_{b_{i_1}}| \geq |A_1| + |B_{b_{i_1}}| - |H_{d_1}| + \rho_1$, for some proper subgroup $H_{d_1} < H_a$, where ρ_1 is the number of H_{d_1} -holes in A_1 and $B_{b_{i_1}}$. Likewise $|A_2 + B_{b_{i_2}}| \geq |A_2| + |B_{b_{i_2}}| - |H_{d_2}| + \rho_2$.

Hence, since $|H_{d_i}| \leq \frac{1}{2}|H_a|$, and since $|H_a| > |B_{b_i}|$ for all i , it follows in view of (31) that

$$\begin{aligned} |A + B| &\geq |A| + |B| - (|A_1| + |A_2| + \sum_{i=1}^l |B_{b_i}|) + (l-1)|H_a| + \sum_{i=1}^r |C_{c_i}| \geq \\ &|A| + |B| - (|A_1| + |A_2| + |B_{i_1}| + |B_{i_2}|) + |H_a| + |A_1 + B_{b_{i_1}}| + |A_2 + B_{b_{i_2}}| \geq \\ &|A| + |B| + |H_a| - |H_{d_1}| - |H_{d_2}| + \rho_1 + \rho_2 \geq |A| + |B|. \end{aligned} \quad (32)$$

Hence, since $|A + B| = |A| + |B|$, it follows that $r = l = 2$, $\rho_1 = \rho_2 = 0$, and $|H_{d_1}| = |H_{d_2}| = \frac{1}{2}|H_a|$, as otherwise the above estimate will be strict.

Since $r = 2$, then $|H_a| = 2$ would imply that $d^\subseteq(A + B, \mathcal{P}) \leq 2$, a contradiction. Therefore $|H_{d_1}| = |H_{d_2}| = \frac{1}{2}|H_a| > 1$. If $H_{d_1} \cap H_{d_2}$ is nontrivial, then in view of $\rho_1 = \rho_2 = 0$, it follows that A , and thus $A + B$, is periodic, a contradiction. Thus it follows that $|H_a| \geq |H_{d_1}||H_{d_2}| = \frac{1}{4}|H_a|^2$. Hence, since $|H_{d_1}| = |H_{d_2}| = \frac{1}{2}|H_a| > 1$, it follows that $|H_a| = 4$. Thus $d^\subseteq(A + B, A + B + H_{d_1}) \leq |H_a| - |H_{d_2}| = 2$, contradicting that $d^\subseteq(A + B, \mathcal{P}) \geq 3$, and completing the case.

Case 2: Next assume that $l = r = 2$, and that w.l.o.g. $\phi_a(a_1 + \{b_1, b_2\}) = \phi_a(b_1 + \{a_1, a_2\}) = \{\phi_a(c_1), \phi_a(c_2)\}$. Hence, since each C_{c_i} is a partially filled H_a -coset, it follows in view of Proposition 2.1 that

$$|A_1| + |B_{b_2}| \leq |H_a| \quad \text{and} \quad |B_{b_1}| + |A_2| \leq |H_a|. \quad (33)$$

Since $\phi_a(b_1 + \{a_1, a_2\}) = \{\phi_a(c_1), \phi_a(c_2)\}$, it follows that $|C_{c_1}| + |C_{c_2}| \geq |A_2| + |B_{b_1}|$, and from Kneser's Theorem applied to $B_{b_1} + A_1$, it follows that equality is possible only if $C_{c_1} \cup C_{c_2} = B_{b_1} + (A_1 \cup A_2)$ is H_b -periodic with $\phi_b(A_1) = 1$ and $H_b \leq H_a$. Thus, since $A + B$ is aperiodic, it follows that equality is possible only if $|A_1| = 1$. The same argument applied to $\phi_a(a_1 + \{b_1, b_2\}) = \{\phi_a(c_1), \phi_a(c_2)\}$ also shows that $|C_{c_1}| + |C_{c_2}| \geq |B_{b_2}| + |A_1|$, with equality possible only if $|B_{b_1}| = 1$. Hence, since $|A + B| \leq |A| + |B|$, it follows in view of (33) and (31), that we must in fact have equality in both the estimates $|C_{c_1}| + |C_{c_2}| \geq |B_{b_2}| + |A_1|$ and $|C_{c_1}| + |C_{c_2}| \geq |A_2| + |B_{b_1}|$, as well as both inequalities in (33). Consequently, $|A_1| = |B_{b_1}| = 1$, $|A_2| = |B_{b_2}| = |H_a| - 1$, and $C_{c_1} \cup C_{c_2} = B_{b_1} + (A_1 \cup A_2)$. Thus in view of Lemma 4.8, it follows that $d^\subseteq(C, \mathcal{QP}_{H_a}) = d^\subseteq(C, \mathcal{QP}) = 1$ for all $C \in \{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$. Hence, since $d^\subseteq(A + B, \mathcal{P}) \geq 3$, it follows that $|H_a| \geq 3$. Thus, since the C_{c_i} are partially filled, and since $C_{c_1} \cup C_{c_2} = B_{b_1} + (A_1 \cup A_2)$, it follows in view of Proposition 2.1 that $\phi_a(a_2 + b_2) \neq \phi_a(c_i)$ for all i , and that $\phi_a(a_2 + b_1) = \phi_a(a_1 + b_2)$. Consequently, letting α be the H_a -hole in A_2 , and letting β be the H_a -hole in B_{b_2} , it follows that $|A \cup \{\alpha\} + B \cup \{\beta\}| = |A + B| + 1 = |A \cup \{\alpha\}| + |B \cup \{\beta\}| - 1$, yielding (16), and completing the proof. \square

Lemma 4.11. *Let A and B be nonempty subsets of a finite abelian group G with $|A + B| = |A| + |B|$, $0 \in A \cap B$, $|A|, |B|, d^\subseteq(A + B, \mathcal{P}) \geq 3$, (A, B) non-extendible, and $\langle A \rangle = G$. If $d^\subseteq(A, \mathcal{QP}) = 1$ and either $|A| \geq 4$ or $|B| \geq 4$, then $d^\subseteq(C, \mathcal{QP}) = 1$ for all $C \in \{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$ and (16) holds.*

Proof. In view of Lemma 4.10, it follows the proof is complete unless $A = A_1 \cup A_0$ with each A_i a subset of an H_a -coset and $|A_1| = |H_a| - 1$, for some nontrivial subgroup H_a . If A_0 is empty, then $\langle A \rangle = G$ implies that $H_a = G$, whence $|A| = |G| - 1$, contradicting $d^\subseteq(A + B, \mathcal{P}) \geq 3$. Therefore we can assume A_0 is nonempty. Let b_1, \dots, b_c be those elements of B that are the unique element from their H_a -coset in B , and let $B' = B_{b'_1} \cup \dots \cup B_{b'_l}$ be an H_a -coset decomposition of the remaining elements of B . Since A is not quasi-periodic it follows that $|A_0| < |H_a|$. Hence $|A| \geq 3$ implies $|H_a| \geq 3$. In view of Lemma 4.8, it follows that B is not quasi-periodic and $\langle B \rangle = G$. We divide the proof into two cases.

Case 1. Suppose $|\phi_a(A) + \phi_a(B)| \geq |\phi_a(A)| + |\phi_a(B)| - 1 = |\phi_a(B)| + 1$. Hence in view of Proposition 2.1 it follows that

$$\begin{aligned} |A + B| &\geq (l + c)|H_a| - c + |A_0| = \\ &(|H_a| + |A_0| - 1) + (l|H_a| + c) + c(|H_a| - 2) - |H_a| + 1 \geq |A| + |B| + \rho' + (c - 1)(|H_a| - 2) - 1, \end{aligned}$$

where ρ' is the number of H_a -holes in B' . Hence $|A + B| \leq |A| + |B|$ implies

$$\rho' + (c - 1)(|H_a| - 2) - 1 \leq 0. \quad (34)$$

Note $c \geq 1$, since otherwise adding the H_a -hole contained in A_1 will, in view of Proposition 2.1, contradict the non-extendibility of A .

Suppose $\rho' = 0$. If B' is empty, then $|B| \geq 3$ implies $c \geq 3$; otherwise, the cases $c \leq 2$ are covered by Lemma 4.10. Thus in all cases we can assume $c \geq 3$, whence (34) implies $2|H_a| \leq 5$, contradicting $|H_a| \geq 3$. So we can assume $\rho' > 0$.

If $\rho' > 1$, then (34) and $|H_a| \geq 3$ imply $c \leq 0$, a contradiction. Therefore we can assume $\rho' = 1$, whence (34) and $|H_a| \geq 3$ imply $c = 1$. Hence if $|\phi_a(B)| > 2$, then the proof is complete in view of $\rho' = 1$ and Lemma 4.10. Otherwise $\rho' = c = 1$ implies that $B = B_1 \cup B_0$ with both B_i nonempty subsets of disjoint H_a -cosets, $|B_1| = |H_a| - 1$, and $|B_0| = 1$. Thus the hypotheses of the lemma and the case are satisfied by interchanging the roles of A and B , whence $|A_0| = 1$ follows by the above argument as well. Since A is not quasi-periodic and since $\langle A \rangle = G$, it follows that \overline{A} is not quasi-periodic. Likewise for B . Hence $d^\subseteq(B, \mathcal{QP}) = d^\subseteq(\overline{B}, \mathcal{QP}) = d^\subseteq(A, \mathcal{QP}) = d^\subseteq(\overline{A}, \mathcal{QP}) = 1$. In view of Proposition 2.1 and $|H_a| \geq 3$, it follows that $|A_1 + B_1| = |H_a|$ and $A_0 + B_1 = A_1 + B_0$, since otherwise $|A + B| \geq 2|H_a| + 1 > |A| + |B|$, a contradiction. Hence $d^\subseteq(A + B, \mathcal{QP}_{H_a}) = d^\subseteq(\overline{A + B}, \mathcal{QP}_{H_a}) = 1$. Thus, since neither $A + B$ nor $\overline{A + B}$ is quasi-periodic (in view of Lemma 4.8), it follows that $d^\subseteq(A + B, \mathcal{QP}) = d^\subseteq(\overline{A + B}, \mathcal{QP}) = 1$. Finally, by letting α be the H_a -hole in A_1 , and by letting β be the H_a -hole in B_1 , it follows in view of $A_1 + B_0 = A_0 + B_1$ and $|A_1 + B_1| = |H_a|$ that $|A \cup \{\alpha\} + B \cup \{\beta\}| = |A + B| + 1 = |A \cup \{\alpha\}| + |B \cup \{\beta\}| - 1$, yielding (16), and completing the case.

Case 2. Suppose $|\phi_a(A) + \phi_a(B)| = |\phi_a(B)|$. Thus from Kneser's Theorem it follows that $\phi_a(B)$ is periodic with maximal subgroup H_b/H_a , and that $\phi_a(A)$ is contained in an H_b/H_a -coset. Hence,

since $\langle A \rangle = G$, it follows that $H_b = G$ and that G/H_a is cyclic generated by $\phi_a(A_1) - \phi_a(A_0)$. Letting ρ' be the number of H_a -holes in B' , it follows in view of Proposition 2.1 that

$$|A + B| \geq (l + c)|H_a| - c = (|H_a| - 1 + |A_0|) + (l|H_a| + c) + (c - 1)(|H_a| - 2) - |A_0| - 1 \geq |A| + |B| - |A_0| + \rho' + (c - 1)(|H_a| - 2) - 1, \quad (35)$$

with equality possible only if $A_1 + b_i + H_a \not\subseteq A + B$ for $i = 1, \dots, c$. Since $d^\subseteq(A + B, \mathcal{P}) \geq 3$, it follows, in view of Proposition 2.1 and the assumption of the case, that $c \geq 3$, with equality possible only if $A_1 + b_i + H_a \not\subseteq A + B$ for $i = 1, \dots, c$.

Suppose $l > 0$. Hence, since $\phi_a(A_1) - \phi_a(A_0)$ generates G/H_a , and since $c \geq 3 > 0$, it follows that there exists b'_j such that $\phi_a(A_0 + b'_j) = \phi_a(A_1 + b_i)$ for some b_i . Thus it follows in view of Proposition 2.1 that either $\rho' \geq |A_0|$, or else $A_1 + b_i + H_a = A_0 + b'_j + H_a \subseteq A + B$. In the former case, it follows in view of (35), $|A + B| \leq |A| + |B|$ and $|H_a| \geq 3$, that $c \leq 2$, a contradiction. In the latter case, it follows that the inequality (35) is strict and that $c \geq 4$. Hence it follows, in view of $|A + B| \leq |A| + |B|$ and $|A_0| \leq |H_a| - 1$, that $2|H_a| \leq 5$, contradicting $|H_a| \geq 3$. So we may assume $l = 0$, whence $c = |B|$.

Suppose $|A_0| \leq |H_a| - 2$. In view of (35), $|A + B| \leq |A| + |B|$, $|H_a| \geq 3$, and $c \geq 3$, it follows that $|A_0| \geq 2|H_a| - 5 \geq |H_a| - 2$. Thus $|A_0| \leq |H_a| - 2$ implies $|H_a| = 3$, whence $|A_0| \leq |H_a| - 2$ implies $|A| = 3$. Thus by hypothesis $c = |B| \geq 4$, whence (35) implies $|A_0| \geq 3|H_a| - 7 \geq |H_a| - 1$, a contradiction. So we can assume $|A_0| = |H_a| - 1$.

If $A_1 + b_i + H_a \subseteq A + B$ for some i , then (35) will be strict and $c \geq 4$, whence $|A_0| \geq 3|H_a| - 6 \geq |H_a|$, a contradiction. Therefore we can assume $A_1 + b_i + H_a \not\subseteq A + B$ for all i . Thus, since $\phi_a(B)$ is G/H_a -periodic with G/H_a cyclic generated by $\phi_a(A_1) - \phi_a(A_0)$, and since $|A_1| = |A_0| = |H_a| - 1$, it follows that we can permute the b_i such that $A_1 + b_i = A_0 + b_j$ for $i \equiv j + 1 \pmod{c}$. Consequently,

$$\begin{aligned} A + B + \{b_1, b_2\} &= \left(\bigcup_{i=1}^c A_1 + b_i \right) + \{b_1, b_2\} = \left(\bigcup_{i=1}^c A_1 + b_i + b_1 \right) \cup \left(\bigcup_{i=1}^c A_1 + b_2 + b_i \right) = \\ &= \left(\bigcup_{i=1}^c A_1 + b_i + b_1 \right) \cup \left(\bigcup_{i=1}^c A_0 + b_1 + b_i \right) = \left(\left(\bigcup_{i=1}^c A_1 + b_i \right) \cup \left(\bigcup_{i=1}^c A_0 + b_i \right) \right) + b_1 = A + B + b_1, \end{aligned}$$

implying from Kneser's Theorem that $A + B$ is periodic, a contradiction. \square

The following lemma will also be used to transfer unpaired structural information from the single set A to a pair containing A , this time in the case when A is an arithmetic progression.

Lemma 4.12. *Let A and B be finite, nonempty subsets of an abelian group G with $|A + B| = |A| + |B|$, $|A| \geq 3$, and $A + B$ aperiodic. If A is an arithmetic progression with difference d , then $h_d(B) = 1$, $h_d(A + B) = h_d(\overline{A + B}) = 0$ and (16) holds.*

Proof. Note that $|A + B| \geq |B| + c(|A| - 1)$, where c is the number of d -components of \overline{B} with length at least $|A| - 1$. Since $A + B$ is aperiodic, it follows that there must be at least one d -component of \overline{B} with length at least $|A| - 1$. Hence either $|A + B| \geq |B| + 2(|A| - 1)$ or else

$|A + B| = |B| + |A| - 1 + h_d(B)$. Since $|A| \geq 3$, and since $|A + B| = |A| + |B|$, it follows that former cannot hold, whence the latter implies $h_d(B) = 1$. Hence, since $|A| \geq 3$, it follows that $h_d(A + B) = 0$, whence $h_d(\overline{A + B}) = 0$ as well. Letting β be the single hole in B and letting $\alpha \in A$, it follows that (16) holds. \square

The following lemma will be one of our main tools for reducing the case $A + B = C$ to a case $A' + B' = C'$. Lemma 4.13 will allow us to conclude the sets $A' = A + \{0, d\}$, $B' = B$ and $C' = C + \{0, d\}$ also satisfy $|A' + B'| = |A'| + |B'|$ (whence induction will be employed), provided $c_d(A) = 2$ for some nonzero d . We note this was (more or less) the main strategy used to prove the prime order case of Theorem 4.1 in [13]. Lemma 4.13 will also be needed for Lemma 4.14.

Lemma 4.13. *Let A and B be nonempty subsets of a finite abelian group G with $|A + B| = |A| + |B|$, $0 \in A \cap B$, $|A| \geq 4$, $|B| \geq 3$, $d^\subseteq(A + B, \mathcal{P}) \geq 3$, (A, B) non-extendible, $\langle A \rangle = G$ and A not quasi-periodic. If $c_d(A) = 2$ for some nonzero d , then either $c_d(B), c_d(A + B) \leq 2$, or else (16) holds and either $d^\subseteq(C, \mathcal{QP}) = 1$ for all $C \in \{A, B, A + B, \overline{A}, \overline{B}, \overline{A + B}\}$, or $d^\subseteq(B, \mathcal{AP}) = 0$.*

Proof. Since A is non-extendible, it follows in view of Proposition 2.4 that $-B + \overline{A + B} = \overline{A}$. Suppose $c_d(B) \geq 3$. Hence since $c_d(A) = c_d(\overline{A}) = 2$, and since $|A + B| = |A| + |B|$, it follows that

$$|(-B + \{0, d\}) + \overline{A + B}| \leq |-B + \{0, d\}| + |\overline{A + B}| - 1. \quad (36)$$

If $\overline{A} + \{0, d\}$ is periodic, then $|A| \geq 3$ and $c_d(\overline{A}) = 2$ imply that A is a union of a nonempty periodic set and at most two elements, whence Lemma 4.10 implies $d^\subseteq(A, \mathcal{QP}) = 1$. Thus Lemma 4.11 completes the proof. Therefore we can assume that $\overline{A} + \{0, d\}$ is aperiodic. Hence Kneser's Theorem and (36) imply $c_d(B) = 3$ and $|(-B + \{0, d\}) + \overline{A + B}| = |-B + \{0, d\}| + |\overline{A + B}| - 1$, whence we can apply KST to the pair $(-B + \{0, d\}, \overline{A + B})$. Let $-B + \{0, d\} = B_1 \cup B_0$ and $\overline{A + B} = C_1 \cup C_0$ be the Kemperman decompositions with common quasi-period H_a .

In view of Lemma 4.8, it follows that $\overline{A + B}$ is not quasi-periodic and that $\langle \gamma - \overline{A + B} \rangle = G$ for $\gamma \in A + B$. Hence $H_a = G$, whence we cannot have type (I), nor as the sumset is aperiodic can we have type (III).

Suppose we have type (II). Hence $\overline{A + B}$ is an arithmetic progression with difference d' , whence Lemma 4.12 applied to $(-B, \overline{A + B})$ implies $h_{d'}(B) = 1$ and $h_{d'}(\overline{A}) = h_{d'}(A) = 0$. However, in view of Proposition 2.4, and Lemma 4.12 applied to $(-A, \overline{A + B})$, it follows that $h_{d'}(A) = 1$, contradicting $h_{d'}(A) = 0$. So we cannot have type (II), and thus must have type (IV). Hence, since $\langle -\gamma + \overline{A + B} \rangle = G$, for $\gamma \in \overline{A + B}$, it follows that $|\overline{A} + \{0, d\}| = |G| - 1$, implying $|A| \leq 3$, a contradiction. So we can assume $c_d(B) \leq 2$.

Applying the above argument with the roles of B and $\overline{A + B}$ interchanged, it follows that either $c_d(\overline{A + B}) = c_d(A + B) \leq 2$, or else B is an arithmetic progression. Hence, since $c_d(B) \leq 2$, it follows that we can assume the later case holds, else the proof is complete. Thus in view of Lemma 4.12 it follows that (16) holds, completing the proof. \square

The next lemma stretches Lemma 4.12 one step further, to handle the case $d^\subseteq(A, \mathcal{AP}) = 1$.

Lemma 4.14. *Let A and B be nonempty subsets of a finite abelian group G with $|A+B| = |A|+|B|$, $0 \in A \cap B$, $|A|, |B|, d^\subseteq(A+B, \mathcal{P}) \geq 3$, and (A, B) non-extendible, $\langle A \rangle = G$ and A not quasi-periodic. If $d^\subseteq(A, \mathcal{AP}_d) = 1$ for some nonzero $d \in G$, and if at most one of $|A|, |B|$ and $|\overline{A+B}|$ is equal to 3, then (16) holds, and one of $h_d(B) \leq 1$, or $h_{d'}(B) = 0$ for some non-zero d' , or $d^\subseteq(C, \mathcal{QP}) = 1$ for all $C \in \{A, B, A+B, \overline{A}, \overline{B}, \overline{A+B}\}$, also holds*

Proof. Since $d^\subseteq(A, \mathcal{AP}_d) = 1$, it follows that A is a subset of an arithmetic progression with difference d and one hole. Hence, $c_d(A) = 2$. Furthermore, since $\langle A \rangle = G$, it follows that $\langle d \rangle = G$. Since B is not quasi-periodic (in view of Lemma 4.8), it follows that $h_d(B) = 0$ implies B is an arithmetic progression, whence Lemma 4.12 completes the proof. So we can assume $h_d(B) > 0$.

By translating and considering $-A$ and $-B$ if necessary, we may w.l.o.g. assume $0, d \in A$, and that 0 is the first term of the minimal arithmetic progression with difference d containing A . Since $\langle d \rangle = G$, let B_1, \dots, B_c be the d -components of B cyclicly ordered according to the direction given by d . Let B'_i be the d -component of \overline{B} located between B_i and B_{i+1} , with indices taken modulo c . Hence G is the disjoint union of the B_i and B'_i . Observe, since 0 and d are the first two terms in A , that for each i either at least $\min\{|A|, |B'_i|\}$ of the holes contained in B'_i are elements of $A+B$, or else $|B_i| = 1$ and at least $\max\{1, \min\{|A| - 1, |B'_i| - 1\}\}$ of the holes contained in B'_i are elements of $A+B$. Hence, since $|B| + 2(|A| - 1) > |A| + |B|$, it follows that $|B'_i| < |A|$ for all but at most one i (say c).

If $|A| = 3$, then $|A+B| = |B| + 3$ implies $c \leq 3$. Hence if $|B'_i| < |A|$ for all i , then $|A+B| \geq |\langle A \rangle| - c \geq |G| - 3$. Thus $|A| = |\overline{A+B}| = 3$, a contradiction. If $|A| \geq 4$, then in view of Lemma 4.13 it follows that $c \leq 2$, else the proof is complete. Hence if $|B'_i| < |A|$ for all i , then $|A+B| \geq |\langle A \rangle| - c \geq |G| - 2$, contradicting $d^\subseteq(A+B, \mathcal{P}) \geq 3$. Thus regardless we can assume $|B'_c| \geq |A|$. Consequently, the discussion of the previous paragraph implies

$$|A+B| \geq |A| + |B| + h_d(B) - x + y, \quad (37)$$

where x is the number of B_i with $|B_i| = 1$, and y is the number of B'_i with $|B_i| = |B'_i| = 1$.

Suppose $|A| = 3$. Hence $A = \{0, d, 3d\}$, and as noted in the previous paragraph, $c \leq 3$. Since $|B| \geq 4$, it follows that $x \leq c - 1 \leq 2$. On the other hand, $h_d(B) \geq c - 1$, with equality possible only if $|B'_i| = 1$ for $i \leq c - 1$. Thus from (37) it follows that $|A| + |B| = |A+B| \geq |A| + |B| + h_d(B) - x + y \geq |A| + |B|$, whence indeed $|B'_i| = 1$ for $i \leq c - 1$ and $y = 0$. Hence if $c = 2$, then $h_d(B) = 1$, whence letting α be the hole in A and letting β be the hole in B yields (16) (recall that $\langle d \rangle = G$). If $c = 1$, then $h_d(B) = 0$, a contradiction to the conclusion of the first paragraph. Therefore assume $c = 3$. Hence, since each B'_i contributes at least one to the sumset, and since B'_3 contributes at least $|A| - 1$ (in view of the discussion in the second paragraph of the proof), it follows that $|A+B| \geq |B| + |A| - 1 + (c - 1) > |A| + |B|$, again a contradiction. So we may assume $|A| \geq 4$, and (as noted before) that $c \leq 2$.

Since $|B| \geq 3$ and since $c \leq 2$, it follows that at most one B_i can have cardinality one. Hence from (37) it follows that $|A + B| \geq |A| + |B| + h_d(B) - 1$, whence $h_d(B) = 1$. Letting α be the hole in A , and letting β be the hole in B yields (16), and completes the proof. \square

We conclude the list of lemmas with a short proof of a special case of the Fainting Lemma from [11]. We remark that the idea for the proof of Lemma 4.15 could be used to prove a weaker form of the Fainting Lemma that does not require the assumption about the first isoperimetric number.

Lemma 4.15. *Let A and B be finite, nonempty subsets of an abelian group G with $0 \in A \cap B$, $|A| = 3$, $|A + B| = |A| + |B| + m$, and $\langle A \rangle = G$. If $|\nu_c(A, B)| \geq 2$ for all $c \in A + B$, then G is finite and $|B| \geq |G| - \binom{m+4}{2}$.*

Proof. In view of Proposition 2.2, it follows that $|\nu_c(B + (i-1)A, A)| \geq 2$ for all $c \in A + iB$ and $i \geq 1$. Hence, since $|A| = 3$, it follows that $N_i^{A^*} = N_i^{\leq A^*} = N_i$ for all $i \geq 1$, where $A^* = A \setminus 0$. Thus in view of Proposition 2.3 it follows that

$$N_i - A^* \subseteq N_{i-1}, \quad (38)$$

for all $i \geq 2$. Note since $\langle A \rangle = G$, that either $B + lA = G$ for sufficiently large l , if G is finite, or else $|B + lA| > |B + (l-1)A|$ for all l (in view of Kneser's Theorem), if G is infinite. Thus if we can show that $|N_i| < |N_{i-1}|$ for nonempty N_{i-1} with $i \geq 2$, it will follow that $N_i = \emptyset$ for sufficiently large i , whence G is finite, and that $|B| = |G| - \sum_{i \geq 1} |N_i| \geq |G| - \sum_{i=0}^{m+2} (m+3-i) = |G| - \binom{m+4}{2}$, completing the proof. However, if $|N_i| \geq |N_{i-1}| > 0$, then in view of (38) and Kneser's Theorem, it follows that $|N_i| = |N_{i-1}|$, that N_i is periodic with maximal period H_a , and that A^* is a subset of an H_a -coset. Since $|\nu_c(B + (i-1)A, A)| \geq 2$ for all $c \in B + iA$, it follows that $B + iA = B + (i-1)A + A^*$. Thus, since $|\phi_a(A^*)| = 1$, it follows that $\phi_a(B + (i-1)A) = \phi_a(B + iA)$, whence $N_i = (B + iA) \setminus (B + (i-1)A)$ cannot be H_a -periodic, a contradiction. Therefore $|N_i| < |N_{i-1}|$, completing the proof. \square

We are now ready to proceed with the proof of Theorem 4.1. However, before beginning, we sketch the main points to outline the strategy. We begin by handling the case $A + B$ periodic, and then show that we can restrict our attention to the case when neither A nor B is quasi-periodic. We then handle the cases when $d^\subseteq(A + B, \mathcal{P})$ is small. The assumption that $d^\subseteq(A + B, \mathcal{P})$ is small will allow us (in most instances) to show (16) fairly easily, since adding H -holes to A or B can only increase $A + B$ by at most $d^\subseteq(A + B, \mathcal{P}_H)$ elements. However, there will be one difficult instance that will instead lead to the type (VIII) pair. Once we have established that $d^\subseteq(A + B, \mathcal{P}) \geq 3$, we can restrict our attention to generating subsets and begin to gain access to the lemmas we have just proved. To gain full access, we must handle the case when $|A| \leq 3$. The case $|A| = |B| = 3$ is handled by brute force. We then restrict our attention to the case G finite. For the case $|A| = 3$ with $|B| \geq 4$, we use Lemma 4.15 to show the existence of a unique expression element $a + b$, and proceed by inductive arguments used on the pair $(A, B \setminus b)$. These will fail if $A \subseteq B$ and $|B| = 4$,

in which case an additional argument is used. With the cases $\min\{|A|, |B|\} \leq 3$ complete, the proof then continues, for G finite, by induction (assuming the theorem true for A' and B' with $\min\{|A'|, |B'|\} < \min\{|A|, |B|\}$ or $\min\{|A'|, |B'|\} = \min\{|A|, |B|\}$ and $|A'| + |B'| > |A| + |B|$). We employ the previously mentioned Dyson e -transform as the method to obtain the pairs $A' + B' = C'$. The arguments from the proof of Kneser's Theorem in [10] will be extended to handle the case when $A(e) + B(e)$ is periodic. The cases $|B(e)| \geq 3$ are handled by applying the induction hypothesis to the pair $(A(e), B(e))$. This method fails when $|B(e)| \leq 2$, since in these cases the unpaired structural information gained for a single set is insufficient to directly transfer back to the original pair. In the case $|B(e)| = 1$, we instead use the method developed in Section 3. In the case $|B(e)| = 2$, then via Proposition 2.4 we will obtain $c_d(A) \leq 2$ for some nonzero d , whence we instead consider $A + \{0, d\} + B$, as discussed before Lemma 4.13. We will encounter problems if $|\overline{A+B}|$ is small. The remaining cases will then be shown to follow from the case $|A| = |B| = 4$ with $c_d(A) = c_d(B) = c_d(A+B) = 2$. The proof with G finite concludes by completing this last remaining case directly. The case when G is infinite is then derived from the finite case by the use of an appropriate *Freiman isomorphism* of (A, B) (which is an injective map $\varphi : A \cup B \rightarrow G'$, with G' an abelian group, such that $\varphi(a_1) + \varphi(b_1) = \varphi(a_2) + \varphi(b_2)$ holds, where $a_i \in A$ and $b_i \in B$, if and only if $a_1 + b_1 = a_2 + b_2$).

Proof. We may assume w.l.o.g. that $0 \in A \cap B$. If either A or B is extendible, then (16) immediately follows. Therefore we can assume otherwise, whence Proposition 2.4 implies

$$-A + \overline{A+B} = \overline{B} \quad \text{and} \quad -B + \overline{A+B} = \overline{A}.$$

Suppose that $A+B$ is periodic with maximal period H_a . If A and B are not both H_a -periodic, then w.l.o.g. there exists $\alpha \in \overline{A}$ such that $\phi_a(A) = \phi_a(A \cup \{\alpha\})$. Hence, since $A+B$ is H_a -periodic, it follows that $A \cup \{\alpha\} + B = A+B$, contradicting that A is non-extendible. Therefore we may assume that A and B are both H_a -periodic. From Kneser's Theorem it follows that

$$|\phi_a(A+B)| \geq |\phi_a(A)| + |\phi_a(B)| - 1.$$

If equality holds in the above inequality, then since A and B are both H_a -periodic, it follows that

$$|A+B| = |H_a| |\phi_a(A+B)| = |H_a + A| + |H_a + B| - |H_a| = |A| + |B| - |H_a| < |A| + |B|,$$

a contradiction. If $|\phi_a(A+B)| = |\phi_a(A)| + |\phi_a(B)|$, then the proof is complete. Otherwise

$$|A+B| = |H_a| |\phi_a(A+B)| \geq |H_a + A| + |H_a + B| + |H_a| = |A| + |B| + |H_a| > |A| + |B|,$$

a contradiction once more. So we may assume that $A+B$ is aperiodic.

Next suppose, for some $\gamma \in \overline{A+B}$, that $A+B \cup \{\gamma\}$ is periodic with maximal period H_a . Note $\phi_a(\gamma) \in \phi_a(A+B)$. Hence choosing $\alpha \in (\gamma - B) \cap (H_a + A)$ and $\beta \in (\gamma - A) \cap (H_a + B)$, it follows that $A \cup \{\alpha\} + B \cup \{\beta\} = A+B \cup \{\gamma\}$, whence (16) holds. So we may assume $d^\subseteq(A+B, \mathcal{P}) \geq 2$.

If $|A| = 1$, then $|A + B| = |A| + |B|$ cannot hold, and if $|A| = 2$, then the theorem holds with type (V) and group G . So we can assume $|A|, |B| \geq 3$.

It is readily checked that the theorem holding for A_0 and B_0 in Lemma 4.6 implies that the theorem holds for A and B . Thus, since the case $A + B$ periodic is complete, it follows in view of Lemma 4.6 (by considering reduced quasi-periodic decompositions) that it suffices to prove the theorem when neither A nor B is quasi-periodic. Thus we henceforth assume this is the case. Hence, since $|A|, |B| \geq 3$, it follows in view of Lemma 4.7 that w.l.o.g. we may assume $\langle A \rangle = \langle B \rangle = G$.

Suppose, for some distinct $\gamma_1, \gamma_2 \in \overline{A+B}$, that $A + B \cup \{\gamma_1, \gamma_2\}$ is periodic with maximal period H_a . Note that $\phi_a(\gamma_i) \in \phi_a(A + B)$, for $i = 1, 2$, since otherwise $A + B$ is periodic, a contradiction. Hence choosing $\alpha_1 \in (\gamma_1 - B) \cap (H_a + A)$, and $\beta_1 \in (\gamma_2 - A) \cap (H_a + B)$, it follows that $A \cup \{\alpha_1\} + B \cup \{\beta_1\} = A + B \cup \{\gamma_1, \gamma_2\}$. Since $d^\subseteq(A + B, \mathcal{P}) \geq 2$, it follows that either A or B , w.l.o.g. A , contains at least two H_a -holes. Thus we can find $\alpha_2 \in (H_a + A) \cap \overline{A \cup \{\alpha\}}$ such that $A \cup \{\alpha_1, \alpha_2\} + B \cup \{\beta_1\} = A + B \cup \{\gamma_1, \gamma_2\}$ is maximally H_a -periodic. Hence from Kneser's Theorem it follows that there are $\rho = |H_a| - 1 + 3 = |H_a| + 2$ holes contained among the sets A and B , and that

$$|\phi_a(A + B)| = |\phi_a(A)| + |\phi_a(B)| - 1. \quad (39)$$

Let $(\phi_a(a_1), \phi_a(b_1)), (\phi_a(a_2), \phi_a(b_2)), \dots, (\phi_a(a_l), \phi_a(b_l)) \in \phi_a(A) \times \phi_a(B)$, with $a_i \in A$ and $b_i \in B$, be those pairs from $\phi_a(A) \times \phi_a(B)$ such that $\phi_a(a_i + b_i) \in \{\phi_a(\gamma_1), \phi_a(\gamma_2)\}$. Since $\gamma_i \notin A + B$, it follows in view of Proposition 2.1 that

$$|A_{a_i}| + |B_{b_i}| \leq |H_a|, \quad (40)$$

for all i .

Suppose $|\{\phi_a(a_i)\}_{i=1}^l| = 1$. Hence in view of the non-extendibility of A , it follows that $a + H_a \subseteq A$ for all $a \in A \setminus A_{a_1}$. Thus, since A is not-quasi-periodic, it follows that $A = A_{a_1}$, whence $\langle A \rangle = G$ implies $H_a = G$. Hence, since there are exactly $|G| - |B|$ elements $\alpha \in G$ such that $\gamma_1 \notin \alpha + B$, and since $|A| + |B| = |G| - 2$, it follows that there exists such an $\alpha \in \overline{A}$. Likewise, since there are exactly $|G| - |A| - 1$ elements $\beta \in G$ such that $\gamma_1 \notin \beta + (A \cup \{\alpha\})$, and since $|A| + |B| = |G| - 2$, it follows that there exists such a $\beta \in \overline{B}$. Hence $A \cup \{\alpha\} + B \cup \{\beta\} \subseteq G \setminus \gamma_1$, whence the non-extendibility of (A, B) implies equality, yielding (16). So we can assume $|\{\phi_a(a_i)\}_{i=1}^l| \geq 2$. By the same argument, it also follows that $|\{\phi_a(b_i)\}_{i=1}^l| \geq 2$.

Hence in view of (40), it follows that $\rho \geq 2|H_a|$, with equality possible only if $|\{\phi_a(a_i)\}_{i=1}^l| = 2$ and $|\{\phi_a(b_i)\}_{i=1}^l| = 2$. Thus in view of $\rho = |H_a| + 2$, it follows that $|H_a| = 2$, that $|\{\phi_a(a_i)\}_{i=1}^l| = 2$ and that $|\{\phi_a(b_i)\}_{i=1}^l| = 2$, implying $\phi_a(\gamma_1) \neq \phi_a(\gamma_2)$ (else $A + B$ is periodic). There are three cases for l .

Suppose $l = 2$. Hence w.l.o.g. $\phi_a(A_{a_1}) + \phi_a(B_{b_1}) = \phi_a(\gamma_1)$ and $\phi_a(A_{a_2}) + \phi_a(B_{b_2}) = \phi_a(\gamma_2)$ are both unique modulo H_a expression elements, whence letting α be the other element from the H_a -coset $a_1 + H_a$, and letting β be the other element from the H_a -coset $b_1 + H_a$, it follows that $A \cup \{\alpha\} + B \cup \{\beta\} = A + B \cup \{\gamma_1\}$, yielding (16). So we can assume $l > 2$

Suppose $l = 3$. Hence it follows that some $\phi_a(a_i)$, say $\phi_a(a_{j_1})$, is contained in only one pair $(\phi_a(a_i), \phi_a(b_i))$. Likewise some $\phi_a(b_i)$, say $\phi_a(b_{j_2})$, is also only contained in one pair $(\phi_a(a_i), \phi_a(b_i))$. Hence, since $l = 3$, it follows that $b_{j_1} \neq b_{j_2}$ and $a_{j_1} \neq a_{j_2}$. Thus neither $\phi_a(a_{j_2} + b_{j_2})$ nor $\phi_a(a_{j_1} + b_{j_1})$ can equal $\phi_a(a_{j_2} + b_{j_1}) \in \{\phi_a(\gamma_1), \phi_a(\gamma_2)\}$, whence $\phi_a(a_{j_1} + b_{j_1}) = \phi_a(a_{j_2} + b_{j_2})$, and w.l.o.g. assume $\phi_a(a_{j_1} + b_{j_1}) = \phi_a(\gamma_1)$. Hence, letting $\alpha \in (\gamma_1 - B) \cap (A_{a_{j_1}} + H_a)$ and letting $\beta \in (\gamma_1 - A) \cap (B_{b_{j_2}} + H_a)$, it follows that $A \cup \{\alpha\} + B \cup \{\beta\} = A + B \cup \{\gamma_1\}$, whence (16) holds. So we can assume $l = 4$.

Let $A + B = C$. Hence, since each C_{γ_i} is a partially filled H_a -coset with $|H_a| = 2$, it follows in view of $l = 4$ that

$$C_{\gamma_1} \cup C_{\gamma_2} = (A_{a_1} \cup A_{a_2}) + (B_{b_1} \cup B_{b_2}) = b_1 + (A_{a_1} \cup A_{a_2}) = a_1 + (B_{b_1} \cup B_{b_2}),$$

implying from Kneser's Theorem that $(A_{a_1} \cup A_{a_2}) + (B_{b_1} \cup B_{b_2}) = b_1 + (A_{a_1} \cup A_{a_2}) = a_1 + (B_{b_1} \cup B_{b_2})$ is periodic with maximal period $H_{a'}$. Since $|B_{b_1} \cup B_{b_2}| = 2$, it follows that $|H_{a'}| = 2$. Let $H_b = H_a \times H_{a'} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Since $C_{\gamma_1} \cup C_{\gamma_2}$ is an $H_{a'}$ -coset it follows that it does not contain an H_a -coset. Hence $\phi_b(\gamma_1) = \phi_b(a_1) + \phi_b(b_1)$ must be a unique expression element in $\phi_b(A + B)$.

In view of the maximality of H_a , it follows that $\phi_a(A + B)$ is aperiodic. Hence, in view of (39), it follows that we can apply KST to the pair $(\phi_a(A), \phi_a(B))$. Let $\phi_a(A'_1) \cup \phi_a(A'_0)$ and $\phi_a(B'_1) \cup \phi_a(B'_0)$ be the Kemperman decompositions with quasi-period H_d/H_a , where $A = A'_1 \cup A'_0$ and $B = B'_1 \cup B'_0$. Since $\phi_b(\{a_1, a_2\} + \{b_1, b_2\})$ is a unique expression element, it follows that both elements in $\phi_a(\{a_1, a_2\} + \{b_1, b_2\})$ have exactly two representations in $\phi_a(A) + \phi_a(B)$ given by

$$\phi_a(a_1) + \phi_a(b_2) = \phi_a(a_2) + \phi_a(b_1) \text{ and } \phi_a(a_1) + \phi_a(b_1) = \phi_a(a_2) + \phi_a(b_2). \quad (41)$$

Since $\phi_a(A + B)$ is aperiodic, it follows that we cannot have type (III). Suppose we have type (II). If say $a_1 \in A'_1$, then (since type (II) implies $|A_x|, |B_y| \geq 2$ for all $x \in A$ and $y \in B$), it follows in view of (41) that $\phi_d(a_1) = \phi_d(a_2)$ and that $\phi_d(b_1) = \phi_d(b_2)$. Furthermore, if $b_i \in B'_1$, then $|H_d/H_a| = 2$, and if $b_i \in B'_0$, then $|\phi_a(B'_0)| = 2$. However, since $\phi_a(B'_0)$ is an arithmetic progression, and since $\{\phi_a(b_1), \phi_a(b_2)\}$ is periodic with period H_b/H_a , it follows that $|H_d/H_a| = 2$ in this latter case as well. However, $|H_d/H_a| = 2$ is impossible for type (II) (since the order of the difference of the arithmetic progression given by KST is at least $|A'_0| + |A'_1| - 1 \geq 3$). Therefore, we can assume $a_1, a_2 \in A'_0$ and $b_1, b_2 \in B'_0$. Hence, since $\phi_a(\{a_1, a_2\})$ is an H_b/H_a -coset with $|H_b/H_a| = 2$, and since $\phi_a(A'_0)$ is an arithmetic progression whose difference generates the cyclic group $\langle \phi_a(A'_0) - \phi_a(a_1) \rangle = \langle \phi_a(B'_0) - \phi_a(b_1) \rangle \leq H_d/H_a$, it follows that $|\phi_a(A'_0)| > \frac{|\langle \phi_a(A'_0) - \phi_a(a_1) \rangle|}{2}$. Likewise $|\phi_a(B'_0)| > \frac{|\langle \phi_a(B'_0) - \phi_a(b_1) \rangle|}{2} = \frac{|\langle \phi_a(A'_0) - \phi_a(a_1) \rangle|}{2}$, whence Proposition 2.1 implies that $\phi_a(A + B)$ is $\langle \phi_a(A'_0) - \phi_a(a_1) \rangle$ -periodic, a contradiction. So type (II) cannot occur.

Suppose we have type (I) with w.l.o.g. $|\phi_a(A'_0)| = 1$. Hence some a_i , say a_2 , is contained in A'_1 . Thus, if some b_i is contained in B'_0 , then it follows in view of (41) that either $|\phi_a(B'_0)| = 2$, $\phi_d(a_1) = \phi_d(a_2)$ and $b_1, b_2 \in B'_0$ (if $|\phi_a(B'_0)| \geq 2$), or else $\phi_a(A'_0) = \{\phi_a(a_1)\}$ and w.l.o.g. $\phi_a(B'_0) = \{\phi_a(b_2)\}$ (if $|\phi_a(B'_0)| = 1$). In the former case $\phi_a(B'_0)$, and thus $\phi_a(B)$, is H_b/H_a -periodic, contradicting that $\phi_a(A + B)$ is aperiodic. In the later case, $\phi_a(a_1) + \phi_a(b_2) = \phi_a(a_2) + \phi_a(b_1)$ contradicts

that $\phi_a(A'_0) + \phi_a(B'_0) = \phi_a(a_1) + \phi_a(b_2)$ is a unique expression element. Therefore we can assume $b_1, b_2 \in B'_1$. Thus, since $a_2 \in A'_1$, it follows in view of (41) that $\phi_d(a_1) = \phi_d(a_2)$, $\phi_d(b_1) = \phi_d(b_2)$, and $|H_d/H_a| = 2$. Since $\phi_d(a_1) = \phi_d(a_2)$, it follows that $H_b/H_a \leq H_d/H_a$. Hence $|H_d/H_a| = 2$ implies that $H_d = H_b$. Since $|H_d/H_a| = 2$, and since $\phi_a(A + B)$ is aperiodic, it follows that $|\phi_a(A'_0)| = |\phi_a(B'_0)| = 1$. Since $H_d = H_b$, and since $\phi_b(a_1) + \phi_b(b_1)$ is a unique expression element, it follows that $\phi_d(a_1) + \phi_d(b_1)$ is a unique expression element in addition to the unique expression element $\phi_d(A'_0) + \phi_d(B'_0)$. Hence, since $a_1 \notin A'_0$ and since $b_1 \notin B'_0$, then applying KST modulo $H_d = H_b$, it follows that we must have type (II) with (by appropriate choice of sign) both $\phi_d(a_i)$ and $\phi_d(b_i)$ the first term in their respective arithmetic progression, and both $\phi_d(A'_0)$ and $\phi_d(B'_0)$ the last term in their respective arithmetic progression. Thus the theorem holds with type (VIII).

Finally, suppose we have type (IV). Hence $|H_d/H_a| \geq 6$, $|\phi_a(A'_0)|, |\phi_a(B'_0)| \geq 3$, and $\phi_a(A'_0)$ is aperiodic (all consequences for a type (IV) pair). Thus it follows, by the same argument used in the case of type (II), that $a_1, a_2 \in A'_0$ and $b_1, b_2 \in B'_0$. Hence $H_b < H_d$. Thus, since $\phi_b(a_i) + \phi_b(b_i)$ is a unique expression element, it follows in view of Proposition 2.1 that

$$|\phi_b(A'_0 + B'_0)| \geq |\phi_b(A'_0)| + |\phi_b(B'_0)| - 1. \quad (42)$$

In view of the description of type (IV), it follows that $|\phi_b(A'_0)| = l_1 + l_p$ and $|\phi_b(B'_0)| = l_2 + l_p$, where l_p is the number of partially filled H_b/H_a -cosets in $\phi_a(A'_0)$, which is also equal to the number of partially filled H_b/H_a -cosets in $\phi_a(B'_0)$, where l_1 is the number of $\phi_a(x) \in \phi_a(A'_0)$ with $\phi_a(x) + H_b/H_a \subseteq \phi_a(A'_0)$, and where $l_1 + l_2 + l_p = |H_d/H_b|$. Thus $|\phi_b(A'_0)| + |\phi_b(B'_0)| - 1 = |H_d/H_b| + l_p - 1$. Hence, since $|\phi_b(A'_0 + B'_0)| \leq |H_d/H_b|$ holds trivially, and since $\phi_a(A'_0)$ is aperiodic, it follows in view of (42) that $l_p = 1$. Let $\phi_a(x) \in \phi_a(A'_0)$ and $\phi_a(y) \in \phi_a(B'_0)$ be the elements that correspond to the unique partially filled H_b/H_a -coset in $\phi_a(A'_0)$ and $\phi_a(B'_0)$, respectively. Hence, since $\phi_a(A + B)$ is aperiodic, and since $l_p = 1$, it follows that $\phi_b(x) + \phi_b(y)$ is a unique expression element in $\phi_b(A'_0 + B'_0)$. However, since $|H_b/H_a| = 2$, it follows that there is only one element from the coset $x + H_b/H_a$ contained in $\phi_a(A'_0)$, and likewise for the coset $y + H_b/H_a$ in $\phi_a(B'_0)$, whence $\phi_a(x) + \phi_a(y)$ is a unique expression element, contradicting that there are no unique expression elements in a type (IV) pair, and completing the case when $d^\subseteq(A + B, \mathcal{P}) \leq 2$. So we can assume $d^\subseteq(A + B, \mathcal{P}) \geq 3$.

Suppose that $|A| = |B| = 3$. If $(A - A) \cap (B - B) = \{0\}$, then $|A + B| = |A||B| = 9 > 6 = |A| + |B|$, a contradiction. Thus w.l.o.g. $A = \{0, d, a_1\}$ and $B = \{0, d, a_2\}$. In view of Lemma 4.12, it follows that we can assume neither A nor B is an arithmetic progression, else the proof is complete. Since A is not quasi-periodic, it follows that no two elements from A can form a coset of an order two subgroup. Likewise for B .

Note

$$A + B = \{0, d, 2d, a_1, a_1 + d, a_2, a_2 + d, a_1 + a_2\}. \quad (43)$$

If $a_1 = a_2$, then the theorem follows with type (VI). If $|\{0, d, 2d\}| \leq 2$, then $\{0, d\}$ is a subgroup of order 2, which we have noted is not the case. Therefore $|\{0, d, 2d\}| = 3$ and $2d \neq 0$. Since neither

A nor B is an arithmetic progression, it follows that $a_1, a_2 \notin \{2d, -d\}$. Hence, since $a_1 + d = a_2$ and $a_2 + d = a_1$ together contradict that $2d \neq 0$, it follows that

$$|\{0, d, 2d, a_1, a_1 + d, a_2, a_2 + d\}| \geq |\{0, d, 2d\}| + |\{a_1, a_1 + d, a_2, a_2 + d\}| \geq 6,$$

with equality possible only if w.l.o.g. $a_2 = a_1 + d$. Thus in view of (43) and $|A + B| = 6$, it follows that

$$a_1 + a_2 = 2a_1 + d \in A + B = \{0, d, 2d, a_1, a_1 + d, a_1 + 2d\},$$

implying that one of the following hold: $2a_1 + d = 0$, $2a_1 = 0$, $2a_1 = d$, $a_1 = -d$, $a_1 = 0$, or $a_1 = d$. The last three equalities are contradictions. If $2a_1 = 0$, then $\{0, a_1\}$ is a coset of an order two subgroup, if $2a_1 = d$, then $A = \{0, d, a_1\} = \{0, 2a_1, a_1\}$ is an arithmetic progression, and if $-2a_1 = d$, then $B = \{0, d, a_1 + d\} = \{0, -2a_1, -a_1\}$ is an arithmetic progression, all contradictions as well. So we can assume w.l.o.g. $|B| \geq 4$.

The Case G Finite. At this point we assume G is finite, and will handle the case G infinite afterwards by a separate argument. Consequently, in view of Lemma 4.8, it follows that neither $A + B$ nor $\overline{A + B}$ is quasi-periodic, and that $\langle \gamma - \overline{A + B} \rangle = G$, where $\gamma \in \overline{A + B}$. In view of Lemma 4.11, it follows that we can assume $d^\subseteq(A, \mathcal{QP})$, $d^\subseteq(B, \mathcal{QP}) \geq 2$, else the proof is complete. In view of Lemma 4.10, and since $|A|, |B|$, $d^\subseteq(A + B, \mathcal{P}) \geq 3$, it follows that we can assume $d^\subseteq(\overline{B}, \mathcal{P})$, $d^\subseteq(\overline{A}, \mathcal{P}) \geq 3$, else the proof is complete, whence in view of Lemma 4.11 and Proposition 2.4, it follows that we can assume $d^\subseteq(\overline{A + B}, \mathcal{QP}) \geq 2$, else $d^\subseteq(B, \mathcal{QP}) = 1$, a contradiction.

Suppose $|N_1^b(A, B)| \geq 2$ for some $b \in B$. If $A + (B \setminus b)$ is periodic with maximal period H_a , then the non-extendibility of B implies that $B \setminus b$ is H_a -periodic, whence B is quasi-periodic, a contradiction. Hence $A + (B \setminus b)$ is aperiodic, whence Kneser's Theorem implies $|N_1^b(A, B)| = 2$. Thus we can apply KST to the pair $(A, B \setminus b)$. Since $\langle A \rangle = G$ and since A is not-quasi-periodic, it follows that the quasi-period from KST must be G . Hence, since $|A|, |B \setminus b| > 1$, and since $|A + (B \setminus b)| \leq |G| - 2$ (in view of $d^\subseteq(A + B, \mathcal{P}) \geq 3$), it follows from KST that both A and $B \setminus b$ are arithmetic progressions, whence Lemma 4.12 completes the proof. So we can assume $|N_1^b(A, B)| \leq 1$ for all $b \in B$. Likewise, $|N_1^a(B, A)| \leq 1$ for all $a \in A$.

If $|A| = |\overline{A + B}| = 3$, then in view of Proposition 2.4 and the completed case $|A| = |B| = 3$, it follows that the theorem holds with type (VII). Therefore we can assume at most one of $|A|, |B|$ and $|\overline{A + B}|$ is equal to three. If $h_d(A) \leq 1$, then $d^\subseteq(A, \mathcal{QP}) \geq 2$ implies that $d^\subseteq(A, \mathcal{AP}) \leq 1$. Likewise for B . Thus in view of Lemmas 4.12 and 4.14, it follows that $h_d(A), h_d(B) \geq 2$ for all nonzero $d \in G$, else the proof is complete. Likewise, if $h_d(\overline{A + B}) \leq 1$, then $d^\subseteq(\overline{A + B}, \mathcal{QP}) \geq 2$ implies that $d^\subseteq(\overline{A + B}, \mathcal{AP}) \leq 1$, whence in view of Proposition 2.4 and Lemmas 4.12 and 4.14, it follows that either $h_d(B) \leq 1$ for some nonzero $d \in G$, or else $d^\subseteq(B, \mathcal{QP}) \leq 1$, both contradictions. Therefore we can assume $h_d(\overline{A + B}) \geq 2$ for all nonzero $d \in G$ as well. Consequently, $d^\subseteq(C, \mathcal{AP}) \geq 2$ for $C \in \{A, B, \overline{A + B}\}$.

Suppose $|A| = 3$. Hence $|B|, |\overline{A + B}| \geq 4$ as noted above. If $|N_1^b(A, B)| = 0$ for all $b \in B$, then Lemma 4.15 implies $|B| \geq |G| - 6$, whence $|\overline{A + B}| \leq 3$, a contradiction. Therefore, since

$N_1^b(A, B) \leq 1$ for all $b \in B$, it follows that we can assume $|N_1^b(A, B)| = 1$ for some $b \in B$. We proceed by induction on $|B|$, with our domain restricted to pairs (A, B') with each member a non-quasi-periodic, generating subset.

Since $|N_1^b(A, B)| = 1$, and since $A+B$ is not quasi-periodic, it follows that $A+(B \setminus b) = (A+B) \setminus \gamma$ is aperiodic with $\gamma \in A+B$. Suppose (16) holds for A and $B \setminus b$. In view of Corollary 4.2, it follows that if the extended pair $A \cup \{\alpha\}$ and $B \setminus b \cup \{\beta\}$ is periodic without a unique expression element, then $A \cup \{\alpha\} + (B \setminus b \cup \{\beta\}) = A + (B \setminus b)$. Thus, since $A + (B \setminus b)$ is aperiodic, it follows that we can apply KST to $A' = A \cup \{\alpha\}$ and $B' = B \setminus b \cup \{\beta\}$. Since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows that $d^\subseteq(A', \mathcal{QP}) \geq 1$. Hence, since $\langle A \rangle = G$, it follows that KST can only hold with subgroup G , whence either $|A+B| \geq |A'+B'| - 1 \geq |G| - 2$, or else $d^\subseteq(A, \mathcal{AP}) \leq 1$, both contradictions. Therefore we can assume (16) does not hold. Consequently, the pair $(A, B \setminus b)$ must be non-extendible. Hence, in view of Lemma 4.8, it follows that $B \setminus b$ is not quasi-periodic and that $\langle B \setminus b \rangle = G$. Thus we can apply the induction hypothesis to the pair $(A, B \setminus b)$ and assume (16) does not hold.

Hence, since $\langle A \rangle = G$ and since A is not quasi-periodic, it follows that $H_a = G$ in Theorem 4.1. Hence, since $|A|, |B \setminus b| \geq 3$, since $|\overline{A + (B \setminus b)}| \geq 4$, and since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows that we must have type (VI) with $H_a = G$, whence $|B| = 4$ and w.l.o.g. $A = \{0, d, x\}$ and $B = \{0, d, x, b\}$ with $N_1^b(A, B) = \{x + b\}$. Since $d^\subseteq(A, \mathcal{AP}) \geq 2$, then it follows that $x \notin \{2d, 3d, -d, -2d\}$, $d \notin \{2x, 3x, -x, -2x\}$ and $0 \notin \{2x - d, 3x - 2d, 2d - x, 3d - 2x\}$. Additionally, A not quasi-periodic implies A does not contain a coset of an order two subgroup. Hence the previous two sentences yield that

$$\{0, d, 2d\}, \{x, x + d\}, \{2x\}$$

are the three distinct d -components of $A + A \subseteq A + B$. Since $|A + B| = 7$, it follows that $b + A$ contains exactly one element distinct from these 6. Since $N_1^b(A, B) = \{x + b\}$, it follows that this element must be $x + b$. Thus the component $\{0, d\} + b$ of $b + A$ must be contained in either $\{0, d, 2d\}$ or $\{x, x + d\}$, whence $b = 0, d$ or x , all contradictions, completing the induction. So we may w.l.o.g. assume $|A| \geq |B| \geq 4$.

Suppose $|\overline{A + B}| \leq 3$. Hence $d^\subseteq(A + B, \mathcal{P}) \geq 3$ implies that $|\overline{A + B}| = 3$. Thus in view of Proposition 2.4, the above case, and Corollary 4.3, it follows that $d^\subseteq(A, \mathcal{QP} \cup \mathcal{AP}) \leq 1$, a contradiction. So we can also assume $|\overline{A + B}| \geq 4$.

If $e + B \subseteq A$ for all $e \in A - B$, then $A - B + B = A$, implying from Kneser's Theorem that A is periodic, a contradiction. Thus we can choose $e \in A - B$ such that $|(e + B) \cap A|$ is maximal subject to $|(e + B) \cap A| < |B|$. Let $B(e) = (e + B) \cap A$, and let $A(e) = (e + B) \cup A$. Note that $|A(e)| + |B(e)| = |A| + |B|$, that $A(e) + B(e) \subseteq e + A + B$, and that $B(e)$ is non-empty. Assume by induction that the theorem holds for non-quasi-periodic, generating subsets A' and B' either with $\min\{|A'|, |B'|\} < \min\{|A|, |B|\} = |B|$, or else with $\min\{|A'|, |B'|\} = \min\{|A|, |B|\} = |B|$ and $|A'| + |B'| > |A| + |B|$. We have already verified the base of the induction when $|A| + |B| \geq |G| - 3$ or when $\min\{|A|, |B|\} \leq 3$.

Case 1: $A(e) + B(e)$ is periodic with maximal period H_a .

Suppose that $B(e)$ is not H_a -periodic. Then $B(e)$ must have an H_a -hole x . Hence, since $A(e) + B(e)$ is H_a -periodic, it follows that

$$(A \cup (e + B)) + (B(e) \cup \{x\}) = A(e) + (B(e) \cup \{x\}) = A(e) + B(e) \subseteq e + A + B.$$

Consequently, $x - e + A \subseteq A + B$ and $x + B \subseteq A + B$. Since x is an H_a -hole in $B(e)$, then either $x \notin A$ or $x - e \notin B$. Hence, in view of the last two sentences, it follows that we can contradict the non-extendibility of the pair (A, B) by either adding x to A (if $x \notin A$) or else by adding $x - e$ to B (if $x - e \notin B$). So we may assume that $B(e)$ is H_a -periodic.

Let ρ be the number of H_a -holes contained in the pair $A(e)$ and $B(e)$, and let ρ' be the number of H_a -holes contained in the pair A and B . Partition the set A into the disjoint sets $A \cap (e + B)$, A_1 and A_2 , where A_1 consists of those elements of A which modulo H_a are contained in $\phi_a(A) \cap \phi_a(e + B)$ but which are not in $A \cap (e + B)$, and where A_2 are the remaining elements of A not contained modulo H_a in $\phi_a(A) \cap \phi_a(e + B)$. Likewise partition the set $e + B = (A \cap (e + B)) \cup B_1 \cup B_2$. Let ρ'' be the number of H_a -holes contained among A_2 and B_2 . Since $A \cap (e + B) = B(e)$ is H_a -periodic, it follows that $\phi_a(A_1) = \phi_a(B_1)$. Hence, since $A_1 \cap B_1$ is empty, it follows that $|A_1| + |B_1| = |A_1 \cup B_1| \leq |H_a| |\phi_a(A_1)|$. Hence, since $A \cap (e + B)$ is H_a -periodic, it follows that

$$\rho = \rho'' + |H_a| \cdot |\phi_a(A_1)| - |A_1| - |B_1| \geq \rho''. \quad (44)$$

Applying Kneser's Theorem to $(A(e), B(e))$, it follows that

$$|A(e) + B(e)| \geq |A| + |B| - |H_a| + \rho. \quad (45)$$

Suppose $b' \in B_2$ with $|\phi_a(B_{b'} + A) \setminus (\phi_a(A(e) + B(e)))| = t \geq 1$. Thus $\rho'' \geq |H_a| - |B_{b'}|$. Hence in view of (45) and (44) it follows that

$$|A + B| \geq |A(e) + B(e)| + t|B_{b'}| \geq |A| + |B| - |H_a| + \rho'' + |B_{b'}| \geq |A| + |B|,$$

whence equality holds. However, equality in the above estimate implies $t = 1$ and $e + A + B = (A(e) + B(e)) \cup (B_{b'} + A)$, whence $A + B$ is quasi-periodic, a contradiction. So we can assume $\phi_a(B_{b'} + A) \subseteq \phi_a(A(e) + B(e))$ for all $b' \in B_2$, whence the non-extendibility of B implies B_2 is H_a -periodic (or empty). By the same argument applied to $a' \in A_2$ it follows that $\phi_a(A_{a'} + e + B) \subseteq \phi_a(A(e) + B(e))$ for all $a' \in A_2$, and that A_2 is H_a -periodic (or empty). Consequently $\rho'' = 0$.

Let $A_1 = A_{\alpha_1} \cup \dots \cup A_{\alpha_n}$ and $B_1 = B_{\beta_1} \cup \dots \cup B_{\beta_n}$ be H_a -coset decompositions of A_1 and B_1 , with $\phi_a(\alpha_i) = \phi_a(\beta_i)$. In view of the result of the previous paragraph and Lemma 4.10, it follows that $n \geq 3$, else the proof is complete.

Note that $e' \in A_{\alpha_i} - B_{\beta_i} \subseteq H_a$ are exactly those elements such that $(e' + B_{\beta_i}) \cap A_{\alpha_i}$ is nonempty. Additionally, since $A \cap (e + B)$ is H_a -periodic, and since $e' \in H_a$, it follows that $A \cap (e + B) \subseteq A \cap (e' + e + B)$. Thus, in view of the previous two sentences, unless $e' + e + B \subseteq A$, then the element $e' + e$ will contradict the maximality of e . Hence in order to avoid this contradiction

we must have: (a) B_2 empty (else w.l.o.g. there will be an H_a -coset $\beta + H_a$ which intersects $e' + e + B$ but not A), and (b) $e' + B_{\beta_i} \subseteq A_{\alpha_i}$ for each $e' \in A_{\alpha_i} - B_{\beta_i}$ (else w.l.o.g. there will be an element from the coset $\alpha_i + H_a$ contained in $e' + e + B$ but not in A'), and (c) $A_{\alpha_i} - B_{\beta_i} = A_{\alpha_j} - B_{\beta_j}$ for all i and j (else w.l.o.g. there will be an element $e' \in A_{\alpha_i} - B_{\beta_i}$ but $e' \notin A_{\alpha_j} - B_{\beta_j}$, whence the elements from the coset $e' + \alpha_j + H_a$ contained in $e' + e + B$ will not be contained in A , but some element from the coset $e' + \alpha_i + H_a$ contained in $e' + e + B$ will be contained in A).

Since $e' + B_{\beta_i} \subseteq A_{\alpha_i}$ for each $e' \in A_{\alpha_i} - B_{\beta_i}$, it follows that $A_{\alpha_i} - B_{\beta_i} + B_{\beta_i} = A_{\alpha_i}$, implying that $B_{\beta_i} - B_{\beta_i} \subseteq H(A_{\alpha_i})$, where A_{α_i} is maximally $H(A_{\alpha_i})$ -periodic. Hence $A_{\alpha_i} - B_{\beta_i} = -\beta_i + A_{\alpha_i}$. Thus, since $A_{\alpha_i} - B_{\beta_i} = A_{\alpha_j} - B_{\beta_j}$ for all i and j , it follows that $A_{\alpha_i} = A_{\alpha_j} + (\beta_i - \beta_j)$. Consequently, the A_{α_i} are all just translates of one another, implying that $H(A_{\alpha_i}) = H(A_{\alpha_j}) = H_{ka} \leq H_a$, and that $|\phi_{ka}(B_{\beta_i})| = 1$ (whence $|B_{\beta_i}| \leq |H_{ka}|$), for all i and j . Note H_{ka} must be a proper subgroup of H_a , else $A_{\alpha_i} \cap B_{\beta_i}$ would be nonempty, a contradiction. Thus, since A_2 is H_a -periodic (or empty), it follows that A is H_{ka} -periodic, whence $|H_{ka}| = 1$. Hence $B_{\beta_i} - B_{\beta_i} \subseteq H(A_{\alpha_i}) = H_{ka}$ implies that $|B_{\beta_i}| = 1$ for all i .

For each partially filled H_a -coset F_i in $A + B$, it follows that there must be at least one pair A_{α_i} and B_{β_j} such that $A_{\alpha_i} + B_{\beta_j} \subseteq F_i$. Since $A + B$ is not quasi-periodic, it follows that there are at least two distinct partially filled H_a -cosets in $A + B$. In view of the non-extendibility of A , it follows that each A_{α_i} must have a $B_{\beta_{\sigma(i)}}$ such that $A_{\alpha_i} + B_{\beta_{\sigma(i)}} \subseteq F_j$ for some j . Likewise for each B_{β_i} . Hence in view of Proposition 4.9 and $n \geq 3$, it follows that there exist distinct i and i' and distinct j and j' such that $A_{\alpha_i} + B_{\beta_j}$ and $A_{\alpha_{i'}} + B_{\beta_{j'}}$ are each disjoint from $A(e) + B(e)$ and $\phi_a(\alpha_i + \beta_j) \neq \phi_a(\alpha_{i'} + \beta_{j'})$. Hence in view of (45) it follows that

$$|A + B| \geq |A(e) + B(e)| + |A_{\alpha_i} + B_{\beta_j}| + |A_{\alpha_{i'}} + B_{\beta_{j'}}| \geq |A| + |B| + \rho + |A_{\alpha_i}| + |A_{\alpha_{i'}}| - |H_a|, \quad (46)$$

with equality possible only if there are exactly two partially filled H_a -cosets in $A + B$.

Since $|B_{\beta_i}| = 1$ for all i , and since $\rho'' = 0$, it follows in view of (44) that $\rho = n(|H_a| - 1) - \sum_{i=1}^n |A_{\alpha_i}| \geq 2(|H_a| - 1) - |A_{\alpha_{i'}}| - |A_{\alpha_i}|$. Thus (46) implies $|A + B| \geq |A| + |B| + |H_a| - 2$, whence $|H_a| = 2$ and equality must hold in (46). Hence there are exactly two partially filled H_a -cosets in $A + B$. Thus, since $|H_a| = 2$ implies each partially filled H_a -coset contains one hole, it follows that $d^{\subseteq}(A + B, \mathcal{P}_{H_a}) \leq 2$, contradicting that $d^{\subseteq}(A + B, \mathcal{P}) \geq 3$, and completing the proof. So we may assume $A(e) + B(e)$ is aperiodic, whence it follows in view of Kneser's Theorem that either $|A(e) + B(e)| = |A(e)| + |B(e)| - 1$ or $|A(e) + B(e)| = |A(e)| + |B(e)|$. We proceed based on $|B(e)|$.

Case 2: $|B(e)| \geq 3$.

Suppose $|A(e) + B(e)| = |A(e)| + |B(e)| - 1$. Thus we can apply KST to $A(e) + B(e) = (e + A + B) \setminus \{\gamma\}$ with $\gamma \in e + A + B$. Hence, since $\langle \gamma' - \overline{A + B} \rangle = G$, for $\gamma' \in \overline{A + B}$, and since $d^{\subseteq}(\overline{A + B}, \mathcal{QP}) \geq 2$, it follows that the quasi-period from KST must be G . Hence from KST it follows that either $d^{\subseteq}(\overline{A + B}, \mathcal{AP}) \leq 1$ or else $|A + B| \geq |G|$, both contradictions. So we can assume $A(e) + B(e) = e + A + B$.

Suppose $(A(e), B(e))$ is extendible. Hence we can apply KST to $A(e) \cup \{\alpha\} + B(e) \cup \{\beta\} = e + A + B$. As in the previous paragraph, the quasi-period from KST must be G , whence either $d^\subseteq(\overline{A+B}, \mathcal{AP}) \leq 0$ or else $|A+B| \geq |G| - 1$, both contradictions. So we can assume $(A(e), B(e))$ is non-extendible. Thus, since $\overline{A+B}$ is not quasi-periodic, and since $\langle \gamma - \overline{A+B} \rangle = G$, it follows in view of Proposition 2.4 and Lemma 4.8 that $A(e)$ and $B(e)$ are both non-quasi-periodic, generating subsets, whence the theorem holds for $A(e)$ and $B(e)$ by induction hypothesis. Hence in view of Corollary 4.3 it follows that $d^\subseteq(\overline{A+B}, \mathcal{QP} \cup \mathcal{AP}) \leq 1$, a contradiction.

Case 3: $|B(e)| = 1$.

Let T be the subset of $A - B$ such that $T + B \subseteq A$. Thus we can apply Theorem 3.1 with $k = 1$, whence Theorem 3.1(ii) implies

$$|T| \geq |A| \frac{|A||B| - |A| - |B|}{(|A| + |B|)(|B| - 1)}. \quad (47)$$

If $|T| \leq 1$, then (47) implies $(|B| - 1)|A|^2 - (2|B| - 1)|A| - |B|(|B| - 1) \leq 0$. This is an increasing function of $|A|$ for $|A| > \frac{2|B|-1}{2|B|-2}$, whence $|A| \geq |B| \geq 4$ implies

$$0 \geq (|B| - 1)|B|^2 - (2|B| - 1)|B| - |B|(|B| - 1) = |B|(|B|^2 - 4|B| + 2),$$

contradicting that $|B| \geq 4$. Therefore we can assume $|T| \geq 2$. If $|T| \leq |A| - |B| - 2$, then (47) implies, in view of $|B| \geq 4$, that $|A| \leq \frac{2|B|-|B|^3-|B|^2}{|B|-2} < 0$, a contradiction. Therefore we can assume

$$|T| \geq |A| - |B| - 1. \quad (48)$$

Let $A' = A \setminus (T + B)$. Hence $A = A' \cup (T + B)$ and $|T + B| = |A| - |A'|$. Let $a' \in A'$ and $a \in A$. If $(a' + B) \cap (a + B)$ is nonempty, then $a' + b' = a + b$ for some $b, b' \in B$. Hence $a' \in (a' - b) + B$ and $a = a' - b + b' \in (a' - b) + B$. Hence, if a' and a are distinct, then $|(a' - b + B) \cap A| \geq 2$, whence $a' - b + B \subseteq A$. Hence $a' - b \in T$ and $a' \in T + B$, a contradiction. Therefore every element $a' + b$, for $a' \in A'$ and $b \in B$, is a unique expression element in $A + B$. Hence

$$|A + B| \geq |A'||B| + |(A \setminus A') + B| \geq |A| + |A'|(|B| - 1) = |A| + |B| + (|A'| - 1)(|B| - 1) - 1.$$

Thus in view of $|A+B| = |A| + |B|$ and $|B| \geq 3$, it follows that $|A'| \leq 1$. Thus since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows that $T + B$ is aperiodic.

Suppose $|\nu_b(B, -B)| \geq 2$ for some nonzero b . Hence $|(b + t + B) \cap (t + B)| \geq 2$ for $t \in T$. Thus, since $t + B \subseteq A$, it follows that $|(b + t + B) \cap A| \geq 2$, whence $(b + t + B) \subseteq A$. Consequently, $b + T + B \subseteq A$, whence the definition of T implies that $\{0, b\} + T = T$. Thus, since b is nonzero, it follows in view of Kneser's Theorem that T is periodic, contradicting that $T + B$ is aperiodic. So we can assume $|\nu_b(B, -B)| \leq 1$ for all nonzero b . Consequently, B is a Sidon set and $|2B| = \frac{|B|(|B|+1)}{2}$.

Since $|A'| \leq 1$, it follows that $T + B = A \setminus \alpha$ for some $\alpha \in G$, whence $T + 2B = (A \setminus \alpha) + B$. Hence, since $|N_1^a(B, A)| \leq 1$ for all $a \in A$ (shown in the second paragraph of the case G finite), it

follows that $T + 2B = (A + B) \setminus \gamma$ for some $\gamma \in G$. Thus $A + B$ not quasi-periodic implies that $T + 2B$ is aperiodic.

If the inequality in (48) is strict, then since $T + 2B$ is aperiodic, it follows from Kneser's Theorem that $|A| + |B| \geq |T + 2B| \geq |T| + |2B| - 1 = |A| - |B| + \frac{|B|(|B|+1)}{2} - 1$, implying $|B| < 4$, a contradiction. Therefore we can assume $|T| = |A| - |B| - 1$. Hence (47) implies that

$$|A| \geq |B|(|B| - 1)(|B| + 1). \quad (49)$$

Note (where M is as defined in Theorem 3.1) that

$$-\frac{M}{|B|} = -|T|(|B| - 1) = (|B| - |A| + 1)(|B| - 1) \equiv (|B| + 1)(|B| - 1) \pmod{|A|}. \quad (50)$$

In view of (49) it follows that $1 \leq (|B| + 1)(|B| - 1) \leq |A|$. Hence in view of (50) it follows that $x = (|B| + 1)(|B| - 1)|B|$ in Theorem 3.1. Hence Proposition 3.1(iii) and $|A + B| = |A| + |B|$ imply

$$|A| + |B| = |A + B| \geq \frac{|A|^2|B|^2(|A||B|^2 - |A||B| + |B|^3 - |B|)}{|A||B|^2(|A||B|^2 - |A||B|)} = \frac{|A||B| - |A| + |B|^2 - 1}{|B| - 1},$$

which implies $|B| \leq 1$, a contradiction.

Case 4: $|B(e)| = 2$.

If $|A(e) + B(e)| = |A(e)| + |B(e)| - 1$, then the arguments from the analogous part of Case 2 complete the proof. Therefore we can assume $|A(e) + B(e)| = |A(e)| + |B(e)|$, $A(e) + B(e) = e + A + B$ and $c_d(A(e)) = 2$, for d equal to the difference of elements in $B(e)$. Thus $A(e) + B(e) = e + A + B$ implies that $c_d(A + B) = c_d(\overline{A + B}) \leq c_d(A(e)) = 2$. If $c_d(\overline{A + B}) = 1$, then $\overline{A + B}$ not quasi-periodic implies $\overline{A + B}$ is an arithmetic progression, a contradiction. Therefore $c_d(\overline{A + B}) = 2$. Thus, in view of Lemma 4.13, $d^\subseteq(B, \mathcal{QP} \cup \mathcal{AP}) \geq 2$, and Proposition 2.4, it follows that $c_d(A), c_d(B) \leq 2$. Furthermore, we must have $c_d(A) = c_d(B) = 2$ by the same reasoning used to establish this for $\overline{A + B}$.

Since $d^\subseteq(A + B, \mathcal{P}) \geq 3$ and $c_d(A + B) = 2$, it follows that $d^\subseteq(A + B + \{0, d\}, \mathcal{P}) \geq 1$. Suppose $d^\subseteq(A + B + \{0, d\}, \mathcal{P}) = 1$. Thus by the arguments used to establish the theorem when $d^\subseteq(A + B, \mathcal{P}) = 1$, it follows that (16) holds for A and $B + \{0, d\}$ with $A \cup \{\alpha\} + (B + \{0, d\}) \cup \{\beta\}$ periodic with maximal period H_a . If $A \cup \{\alpha\} + (B + \{0, d\}) \cup \{\beta\}$ contains no unique expression element, then Corollary 4.2 implies that $A \cup \{\alpha\} + (B + \{0, d\}) \cup \{\beta\} = A + B + \{0, d\}$. Hence, since $d^\subseteq(A + B + \{0, d\}, \mathcal{P}) \geq 1$, it follows that we can assume this does not happen, whence we can apply KST to $A \cup \{\alpha\} + (B + \{0, d\}) \cup \{\beta\}$. Since $d^\subseteq(A, \mathcal{QP}) \geq 2$, and since $\langle A \rangle = G$, it follows that the quasi-period from KST must be G , whence $A \cup \{\alpha\} + (B + \{0, d\}) \cup \{\beta\}$ periodic and $c_d(A + B) = 2$ imply that $|\overline{A + B}| \leq 3$, a contradiction. So we can assume $d^\subseteq(A + B + \{0, d\}, \mathcal{P}) \geq 2$.

If the pair $(A, B + \{0, d\})$ is extendible, then since $d^\subseteq(A, \mathcal{QP}) \geq 2$, since $\langle A \rangle = G$, and since $d^\subseteq(A + B + \{0, d\}, \mathcal{P}) \geq 2$, it follows in view of KST that $d^\subseteq(A, \mathcal{AP}) \leq 1$, a contradiction. Therefore we can assume $(A, B + \{0, d\})$ is non-extendible. Thus, since A is not quasi-periodic and since $\langle A \rangle = G$, it follows in view of Lemma 4.6 that $B + \{0, d\}$ is not quasi-periodic. Also, $\langle B + \{0, d\} \rangle \geq \langle B \rangle = G$

implies $\langle B + \{0, d\} \rangle = G$. Thus we can apply the induction hypothesis to the pair $(A, B + \{0, d\})$. Since $|A|, |B + \{0, d\}| \geq 4$, with neither A nor $B + \{0, d\}$ quasi-periodic, it follows that we cannot have type (V-VII). Since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows that we cannot have type (VIII). Thus (16) holds for A and $B + \{0, d\}$. Since $\langle A \rangle = G$, and since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows that the quasi-period from KST must be G . Hence, since $d^\subseteq(A, \mathcal{AP}) \geq 2$, and since $c_d(A + B) = 2$, it follows in view of KST that $|\overline{A + B}| \leq 4$, whence $|\overline{A + B}| = 4$.

Suppose $|B| \geq 5$. Hence we can apply the induction hypothesis to $(-A, \overline{A + B})$. Thus, since $d^\subseteq(A, \mathcal{QP}) \geq 2$, it follows in view of Corollary 4.3 that $d^\subseteq(-A, \mathcal{AP}) = d^\subseteq(A, \mathcal{AP}) \leq 1$, a contradiction. So we can assume $|B| = |\overline{A + B}| = 4$. Note that if the theorem holds for $(-B, \overline{A + B})$, then in view of $d^\subseteq(A, \mathcal{QP}) \geq 2$ it follows from Corollary 4.3 that $d^\subseteq(-B, \mathcal{AP}) = d^\subseteq(B, \mathcal{AP}) \leq 1$, a contradiction. Consequently, it follows that case G finite will be complete once we complete the case with $|A| = |B| = 4$ and $c_d(A) = c_d(B) = c_d(A + B) = 2$. We proceed to do so.

Suppose $|N_1^b(A, B)| > 0$ for some $b \in B$. Hence in view of $|N_1^b(A, B)| \leq 1$, it follows that $|N_1^b(A, B)| = 1$. If the pair $(A, B \setminus b)$ is extendible, then the theorem holds for the pair $(A, B \setminus b)$. Otherwise, since A is not quasi-periodic, since $\langle A \rangle = G$, and since $A + B$ is not quasi-periodic (implying $A + (B \setminus b)$ is aperiodic), it follows from Lemma 4.8 that $B \setminus b$ is not quasi-periodic and $\langle B \setminus b \rangle = G$, whence the theorem holds for the pair $(A, B \setminus b)$ by induction hypothesis. In the latter case, Corollary 4.3 implies $d^\subseteq(A, \mathcal{AP} \cup \mathcal{QP}) \leq 1$, a contradiction. In the former case, we can apply KST to $A \cup \{\alpha\}$ and $B \cup \{\beta\}$. Since $d^\subseteq(A, \mathcal{QP}) \geq 2$, and since $\langle A \rangle = G$, it follows that the quasi-period must be G , whence $|\overline{A + B}| \geq 4$ likewise implies that $d^\subseteq(A, \mathcal{AP}) \leq 1$, again a contradiction. So we can assume there are no unique expression elements in $A + B$.

Let A_1 and A_2 be the two d -components of A , and let B_1 and B_2 be the two d -components of B . Suppose for some i , say $i = 1$, that $|A_1| \geq 3$. Hence in view of $h_d(B) \geq 2$, it follows that $|A_1 + B| = |B| + 4 = |B| + |A|$. Thus $A_1 + B = A + B$. Since $c_d(A + B) = 2$, this implies that $A_1 + B_1$ and $A_1 + B_2$ are distinct components in $A + B$, and that each of the four end terms of components in $A_1 + B$ is a unique expression element. We may assume w.l.o.g. that $|B_1| \geq |B_2|$. Thus the component $A_1 + B_1$ is longer than either of the components $A_2 + B_1$ and $A_2 + B_2$. Hence, since $A + B$ contains no unique expression element, and since $A_1 + B = A + B$, it follows that the only way the two end terms of $A_1 + B_1$ will not be unique expression elements in $A + B$ is if $A_2 + B \subseteq A_1 + B_1$. Thus both the end terms of $A_1 + B_2$ are unique expression elements, a contradiction. So we can assume $|A_i| = 2$ for $i = 1, 2$. Likewise $|B_i| = 2$ for $i = 1, 2$.

Since $h_d(B) \geq 2$, it follows that $A_1 + B_1$ and $A_1 + B_2$ are distinct components in $A_1 + B$, and that all four of the end terms are unique expression elements in $A_1 + B$. Since $|A_i + B_j| = |A_{i'} + B_{j'}|$ for all i, i', j, j' , and since $c_d(A + B) = 2$, it follows that only way that these four terms can all not be unique expression elements in $A + B$ is if $A_2 + B_1 = A_1 + B_2$ and $A_2 + B_2 = A_1 + B_1$. Thus $A_1 + B = A_2 + B$, implying $|A + B| = |B| + 2 < |A| + |B|$, a contradiction. Consequently, we conclude that Theorem 4.1 holds for G finite.

The Case G Infinite. Assume G is infinite. Since A and B are finite, we may w.l.o.g. assume

G is finitely generated. Hence $G \cong \mathbb{Z}^l \times T$, where T is the torsion subgroup of G . By translation, we can assume all non-torsion coordinates for all $a \in A$ and $b \in B$ are non-negative. Let M be the maximum integer that occurs in a non-torsion coordinate of the $a \in A$ and $b \in B$. Let p be a prime such that $p > 4(M + |T| + |A||B|)$. Let $\varphi : G \cong \mathbb{Z}^l \times T \rightarrow (\mathbb{Z}/p\mathbb{Z})^l \times T$ be the map defined by reducing all non-torsion coordinates modulo p . Since $p > 4(M + |T| + |A||B|) \geq 2M$, it follows that φ is a Freiman isomorphism of $(A + T, B + T)$ (see the definition given before the start of the proof of Theorem 4.1), and thus also of (A, B) . Hence $|\varphi(A + B)| = |\varphi(A)| + |\varphi(B)| = |A| + |B|$. Thus, since $\varphi(G)$ is finite, it follows that we can apply Theorem 4.1 to $\varphi(A)$ and $\varphi(B)$.

If $d^\subseteq(\varphi(A + B), \mathcal{P}_{H_a}) \leq 2$, for some nontrivial subgroup $H_a \leq \varphi(G)$, then

$$|A| + |B| = |\varphi(A + B)| \geq |H_a| - 2.$$

Hence $p > 4(M + |T| + |A||B|) \geq 4(|A||B| + 1)$ implies $H_a \leq T$ (since any element outside T has a coordinate with order at least p , and thus is itself of order at least p). Hence from the definition of φ it follows that $d^\subseteq(A + B, \mathcal{P}_{H_a}) \leq 2$, a contradiction. Therefore we can assume $d^\subseteq(\varphi(A + B), \mathcal{P}) \geq 3$. Since $\langle A \rangle = G$, it follows that $\langle \varphi(A) \rangle = \varphi(G)$. Likewise $\langle \varphi(B) \rangle = \varphi(G)$.

Suppose that $d^\subseteq(\varphi(A), \mathcal{Q}\mathcal{P}_{H_a}) = 0$ for some nontrivial subgroup $H_a \leq \varphi(G)$. Hence $|\varphi(A)| = |A| \geq |H_a|$. Thus, since $p > 4(M + |T| + |A||B|) \geq 4(|A||B| + 1)$, it follows that $H_a \leq T \leq G$. Hence $d^\subseteq(\varphi(A), \mathcal{Q}\mathcal{P}_{H_a}) = 0$ implies (in view of the definition of φ) that A is quasi-periodic with quasi-period $H_a \leq G$, a contradiction. So we can assume $d^\subseteq(\varphi(A), \mathcal{Q}\mathcal{P}) \geq 1$. Likewise $d^\subseteq(\varphi(B), \mathcal{Q}\mathcal{P}) \geq 1$.

Suppose

$$d^\subseteq(\varphi(A), \mathcal{A}\mathcal{P}_d), d^\subseteq(\varphi(B), \mathcal{A}\mathcal{P}_d) \leq 1, \quad (51)$$

for some nonzero $d \in \varphi(G)$. Consequently, $\varphi(G) = \langle \varphi(A) \rangle$ is cyclic, implying $G \cong \mathbb{Z} \times T$ with T cyclic. For $x \in \varphi(G)$, let \bar{x} denote the least non-negative integer representative of the integer coordinate of x .

If $d \in T$, then it follows that $\varphi(A) \subseteq T$, whence $A \subseteq T$, contradicting that $\langle A \rangle = G$ is infinite. Therefore we can assume $d \notin T$. Hence, by considering $-d$ if needed, it follows that $1 \leq \bar{d} \leq \frac{p-1}{2}$. Since $d^\subseteq(\varphi(A), \mathcal{A}\mathcal{P}_d) \leq 1$, let $P = \{p_0, p_0 + d, \dots, p_0 + |A|d\}$ be an arithmetic progression with difference d that contains $\varphi(A)$.

Suppose $\overline{p_0 + id} = \overline{p_0} + i\bar{d}$ for all i does not hold. Hence, if $d \leq M$, then $p > 4(M + |T| + |A||B|) > 4M$ implies that $\{\overline{p_0}, \overline{p_0 + d}, \dots, \overline{p_0 + |A|d}\}$ must contain at least two elements from the interval (M, p) . Hence P contains at least two elements from \overline{A} , contradicting the $|P \setminus A| \leq 1$. Otherwise, $d \leq \frac{p-1}{2}$ and $M \leq \frac{p-1}{2}$ imply that $\{\overline{p_0}, \overline{p_0 + d}, \dots, \overline{p_0 + |A|d}\}$ contains at least $|A| - 1$ elements from the interval (M, p) , whence $|P \setminus A| \leq 1$ implies $|A| \leq 2$, a contradiction. So we may assume $\overline{p_0 + id} = \overline{p_0} + i\bar{d}$ for all i .

Hence A is contained in an arithmetic progression of difference (\bar{d}, t) and at most one hole, where t is the torsion coordinate of d . By the same argument applied to B , it follows that B is also contained in an arithmetic progression of difference (\bar{d}, t) with at most one hole. Hence letting α

be the hole in A , and letting β be the hole in B , it follows in view of $|A + B| = |A| + |B|$ that (16) holds, completing the proof. So we may assume that (51) does not hold.

Suppose that $(\varphi(A), \varphi(B))$ is extendible. Hence w.l.o.g. there exists $\alpha \in \overline{\varphi(A)}$ such that $\varphi(A) \cup \{\alpha\} + \varphi(B) = \varphi(A) + \varphi(B)$. Thus we can apply KST to $(\varphi(A) \cup \{\alpha\}, \varphi(B))$. Since $d^\subseteq(\varphi(B), \mathcal{QP}) \geq 1$, and since $\langle \varphi(B) \rangle = \varphi(G)$, it follows that the quasi-period from KST must be $\varphi(G)$. Hence, since $d^\subseteq(\varphi(A + B), \mathcal{P}) \geq 3$, it follows from KST that (51) holds, a contradiction. So we can assume $(\varphi(A), \varphi(B))$ is non-extendible.

Since $(\varphi(A), \varphi(B))$ is non-extendible, since $|\overline{\varphi(A + B)}|, |\varphi(B)| \geq 4$, and since (51) does not hold, it follows in view of Corollary 4.3 that $d^\subseteq(\varphi(A), \mathcal{QP}_{H_a}) = 1$ for some nontrivial subgroup $H_a \leq \varphi(G)$. Hence $|\varphi(A)| = |A| \geq |H_a| - 1$. Thus, since $p > 4(M + |T| + |A||B|) \geq 4(|A||B| + 1)$, it follows that $H_a \leq T \leq G$. Observe that we have verified all the hypotheses needed to apply Lemma 4.11 to $(\varphi(A), \varphi(B))$. Hence Lemma 4.11 implies that (16) holds for $(\varphi(A), \varphi(B))$. Thus, since φ is a Freiman isomorphism for $(A + T, B + T)$, since $H_a \leq T$, and since the proof of Lemma 4.11 shows $\alpha \in \varphi(A) + H_a$ and $\beta \in \varphi(B) + H_a$, it follows that (16) holds for (A, B) , completing the proof. \square

Acknowledgements: I wish to thank Oriol Serra for some helpful comments regarding the manuscript.

References

- [1] Y. F. Bilu, V. F. Lev, and I. Z. Ruzsa, Rectification principles in additive number theory, Dedicated to the memory of Paul Erdős, *Discrete Comput. Geom.*, 19 (1998), no. 3, Special Issue, 343–353.
- [2] S. Chowla, A theorem on the addition of residue classes: applications to the number $\Gamma(k)$ in Warings problem, *Proc. Indian Acad. Sc.*, 2 (1935), 242–243.
- [3] J. Deshouillers and G. A. Freiman, A step beyond Kneser’s theorem for abelian finite groups, *Proc. London Math. Soc.*, (3) 86 (2003), no. 1, 1–28.
- [4] G. A. Freiman, Inverse problems of additive number theory, On the addition of sets of residues with respect to a prime modulus, *Doklady Akad. Nauk SSSR*, 141 (1961), 571–573.
- [5] G. A. Freiman, Inverse problems of additive number theory, On the addition of sets of residues with respect to a prime modulus, *Soviet Math.-Doklady*, 2 (1961), 1520–1522.
- [6] G. A. Freiman, *Foundations of a structural theory of set addition*, translated from the Russian, Translations of Mathematical Monographs, Vol 37, American Mathematical Society, Providence, RI, 1973.
- [7] W. T. Gowers, A new proof of Szemerédi’s theorem for arithmetic progressions of length four, *Geom. Funct. Anal.*, 8 (1998), no. 3, 529–551.
- [8] D. J. Grynkiewicz, An extension of the Erdős-Ginzburg-Ziv Theorem to hypergraphs, *European J. Combin.*, 26 (2005), no. 8, 1154–1176.

- [9] D. J. Grynkiewicz, Quasi-periodic decompositions and the Kemperman structure theorem, *European J. Combin.*, 26 (2005), no. 5, 559–575.
- [10] D. J. Grynkiewicz, *Sumsets, zero-sums and extremal combinatorics*, Ph.D. Dissertation, Caltech (2006).
- [11] Y. O. Hamidoune, O. Serra, and G. Zemor, Beyond Chowla’s Theorem, preprint.
- [12] Y. O. Hamidoune, O. Serra, and G. Zemor, On the critical pair theory in $\mathbb{Z}/p\mathbb{Z}$, *Acta Arith.*, 121 (2006), no. 2.
- [13] Y. O. Hamidoune and J. O. Rødseth, An inverse theorem mod p , *Acta Arith.*, 92 (2000), no. 3, 251–262.
- [14] Y. O. Hamidoune, Subsets with a small sum II: The critical pair problem, *European J. Combin.* 21 (2000), no. 2, 231–239.
- [15] Y. O. Hamidoune, Subsets with small sums in abelian groups I: The Vosper property, *European J. Combin.* 18 (1997), no. 5, 541–556.
- [16] Y. O. Hamidoune, A. S. Lladó, and O. Serra, Vosperian and superconnected abelian Cayley digraphs, *Graphs Combin.*, 7 (1991), no. 2, 143–152.
- [17] G. Károlyi, An inverse theorem for the restricted set addition in abelian groups, *J. Algebra*, 290 (2005), no. 2, 557–593.
- [18] G. Károlyi, The Erdős-Heilbronn problem in abelian groups, *Israel J. Math.*, 139 (2004), 349–359.
- [19] J. H. B. Kemperman, On small sumsets in an abelian group, *Acta Math.*, 103 (1960), 63–88.
- [20] M. Kneser, Summenmengen in lokalkompakten abelschen gruppen, *Math. Z.*, 66 (1956), 88–110.
- [21] M. Kneser, Ein satz über abelsche gruppen mit anwendungen auf die geometrie der zahlen, *Math. Z.*, 64 (1955), 429–434.
- [22] M. Kneser, Abschätzung der asymptotischen dichte von summenmengen, *Math. Z.*, 58 (1953), 459–484.
- [23] V. Lev, On small sumsets in abelian groups, *Structure theory of set addition. Asterisque*, No. 258 (1999), xv, 317321.
- [24] V. Lev, Critical Pairs in abelian groups and Kemperman’s Structure Theorem, preprint.
- [25] M. B. Nathanson, *Additive Number Theory: Inverse Problems and the Geometry of Sumsets*, Graduate Texts in Mathematics, 165, Springer-Verlag, New York, 1996.
- [26] O. Serra and G. Zémor, On a generalization of a theorem by Vosper, *Integers* (2000), electronic.
- [27] O. Serra, An isoperimetric method for the small sumset problem, Surveys in combinatorics 2005, *London Math. Soc. Lecture Note Ser.*, 327, Cambridge Univ. Press, Cambridge (2005), 119–152.
- [28] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression, *Acta Math. Acad. Sci. Hungar.*, 20 (1969), 89–104.
- [29] A. G. Vosper, The critical pairs of subsets of a group of prime order, *J. London Math. Soc.*, 31 (1956), 200–205.
- [30] A. G. Vosper, Addendum to “The critical pairs of subsets of a group of prime order,” *J. London Math. Soc.*, 31 (1956), 280–282.